US005915027A

United States Patent [19]

Cox et al.

[11] Patent Number: 5,915,027

[45] Date of Patent: Jun. 22, 1999

[54] DIGITAL WATERMARKING

[75] Inventors: **Ingemar J. Cox**, Lawrenceville, N.J.; **Matthew L. Miller**, Vilnius, Lithuania; **Kazuyoshi Tanaka**; **Yutaka Wakasu**, both of Tokyo, Japan

[73] Assignees: **NEC Research Institute**, Princeton, N.J.; **NEC Corporation**, Tokyo, Japan

[21] Appl. No.: **08/746,022**

[22] Filed: **Nov. 5, 1996**

[51] **Int. Cl.**[6] **H04L 9/02**

[52] **U.S. Cl.** **380/54**

[58] **Field of Search** 388/28, 51, 54

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,939,515	7/1990	Adelson	341/51
5,319,735	6/1994	Preuss et al.	395/2.14
5,530,751	6/1996	Morris	380/4
5,530,759	6/1996	Braudaway et al.	380/54
5,568,570	10/1996	Rabbani	382/238
5,613,004	3/1997	Cooperman et al.	380/28
5,636,292	6/1997	Rhoads	382/232
5,646,997	7/1997	Barton	380/23
5,659,726	8/1997	Sanford, II	395/612
5,687,236	11/1997	Moskowitz et al.	380/28
5,734,752	3/1998	Knox	380/54

FOREIGN PATENT DOCUMENTS

0690595	1/1995	European Pat. Off. .	
2196167	4/1988	United Kingdom .	
8908915	9/1989	WIPO .	
9520291	7/1995	WIPO .	
9621290	7/1996	WIPO	H04H 1/00
9625005	8/1996	WIPO	H04H 7/08
9627259	9/1996	WIPO .	

OTHER PUBLICATIONS

R.G. Van Schyndel et al, "A digital watermark," in Intl. Conf. On Image Processing, vol. 2, pp. 86–90, 1994.

G. Caronni, "Assuring Ownership Rights for Digital Images," in Proc. Reliable IT Systems, VIS '95, 1995.

J. Brassil et al, "Electronic Marking and Identification Techniques to Discourage Document Copying," in Proc. Infocom '94, pp. 1278–1287, 1994.

K. Tanaka et al, "Embedding Secret Information into a Dithered Multi-level Image," in IEEE Military Comm. Conf., pp. 216–220, 1990.

K. Mitsui et al, "Video–Steganography: How to Secretly Embed a Signature in a Picture," in IMA Intellectual Property Project Proc., vol. 1, pp. 187–206, 1994.

Macq and Quisquater, "Cryptology for Digital TV Broadcasting," in Proc. of the IEEE, vol. 83, No. 6, pp. 944–957, 1995.

W. Bender et al, "Techniques for data hiding," in Proc. of SPIE, vol. 2420, No. 40, Jul. 1995.

Koch, Rindfrey and Zhao, "Copyright Protection for Multimedia Data," in Proc. of the Int'l Conf. on Digital Media and Electronic Publishing (Leeds, UK, Dec., 6–8 1994).

Koch and Zhao, "Towards Robust and Hidden Image Copyright Labeling," in Proc. of 1995 IEEE Workshop on Nonlinear Signal and Image Processing (Neos Marmaras, Halkidiki, Greece, Jun. 20–22, 1995.

Zhao and Koch, "Embedding Robust Labels Into Images For Copyright Protection," in Proc. Int. Congr. on IPR for Specialized Information, Knowledge and New Technologies (Vienna, Austria), Aug. 21–25, 1995.

"Digital Copyright: Who Owns What?" NewMedia, Sep. 1995, pp. 38–43.

"Publish and Be Robbed?" New Scientist, Feb. 18, 1995, pp. 32–37.

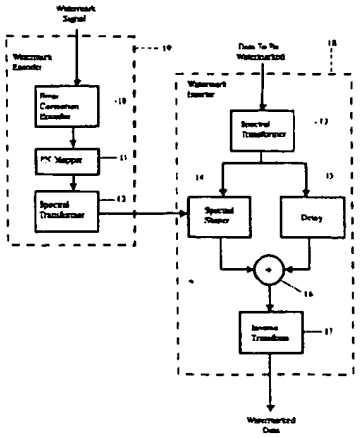
(List continued on next page.)

Primary Examiner—Salvatore Cangialosi
Attorney, Agent, or Firm—Philip J. Feig

[57] **ABSTRACT**

Digital watermarking of data, including image, video and audio data, is performed by repeatedly inserting the watermark into subregions or subimages of the data. Similarly, the watermark is repeatedly extracted from the subregions of the data.

28 Claims, 8 Drawing Sheets



OTHER PUBLICATIONS

Kohno et al, "Spread Spectrum Access Methods for Wireless Communications," in IEEE Communications Magazine, Jan. 1995, pp. 58–67, 116.

Campana and Quinn, "Spread spectrum communications," in IEEE Potentials, Apr. 1993, pp. 13–16.

Mowbray and Grant, "Wideband coding for uncoordinated multiple access communication," in Electronics & Communication Engineering Journal, Dec. 1992, pp. 351–361.

Digimarc Overview & "Wired" Magazine article (Jul. 1995 issue)—(Jun. 1995).

A.G. Bors et al., "Image Watermarking Using DCT Domain Constraints", Dept. Of Informatics, University of Thessaloniki.

I.J. Cox et al., "Secure Spread Spectrum Watermarking for Multimedia", NEC Research Institute, Technical Report 95–10.

H.S. Stone, "Analysis of Attacks on Image Watermarks with Randomized Coefficients", NEC Research Institute, May 17, 1996.

F.M. Boland et al., "Watermarking Digital Images for Copyright Protection", Image Processing and its Applications, Jul. 4–6, 1995, Conference Publication No. 410, pp. 326–330.

L. Boney et al., "Digital Watermarks for Audio Signals".

Swanson et al., "Transparent Robust Image Watermarking", Proc. IEEE Int. Conf. On Image Proc. 1996.

J.J.K. O Ruanaidh et al., "Phase Watermarking of Digital Images".

I. Pitas, "A Method for Signature Casting on Digital Images".

C.T. Hsu et al., "Hidden Signatures in Images", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

M. Schneider et al., "A Robust Content Based Digital Signature for Image Authentication", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

S. Roche et al., "Multi–Resolution Access Control Algorithm Based on Fractal Coding", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

K. Hirotsugu, "An Image Digital Signature System with ZKIP for the Graph Isomorphism", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

R.B. Wolfgang et al., "A Watermark for Digital Images".

J.J.K. O Ruanaidh et al., "Watermarking Digital Images for Copyright Protection", EVA 96 Florence, pp. 1–7.

T. Aura, "Invisible Communication", Nov. 6, 1995.

D. Kahn, "Information Hiding—An Annotated Bibliography", Macmillan 1967, Library of Congress catalog No. 63–16109.

Craver et al., "Can Invisible Watermarks Resolve Rightful Ownerships?", IBM Research Report.

Podilchuk et al., "Digital Image Watermarking Using Visual Models", Proc. of EI '97, vol. 3016, Feb. 9–14, 1997.

Cox et al., "A review of watermarking and the importance of perceptual modeling", proc. of EI'97, vol. 3016, Feb. 9–14, 1997.

Watson, "DCT quantization matrices visually optimized for individual images", SPIE, vol. 1913, pp. 202–216.

Ahumada, Jr. et al., "Luminance–Model–Based DCT Quantization for Color Image Compression", SPIE, vol. 1666 (1992), pp. 365–374.

Hartung et al., "Digital Watermarking of Raw and Compressed Video", Systems for Video Communication, Oct. 1996, pp. 205–213.

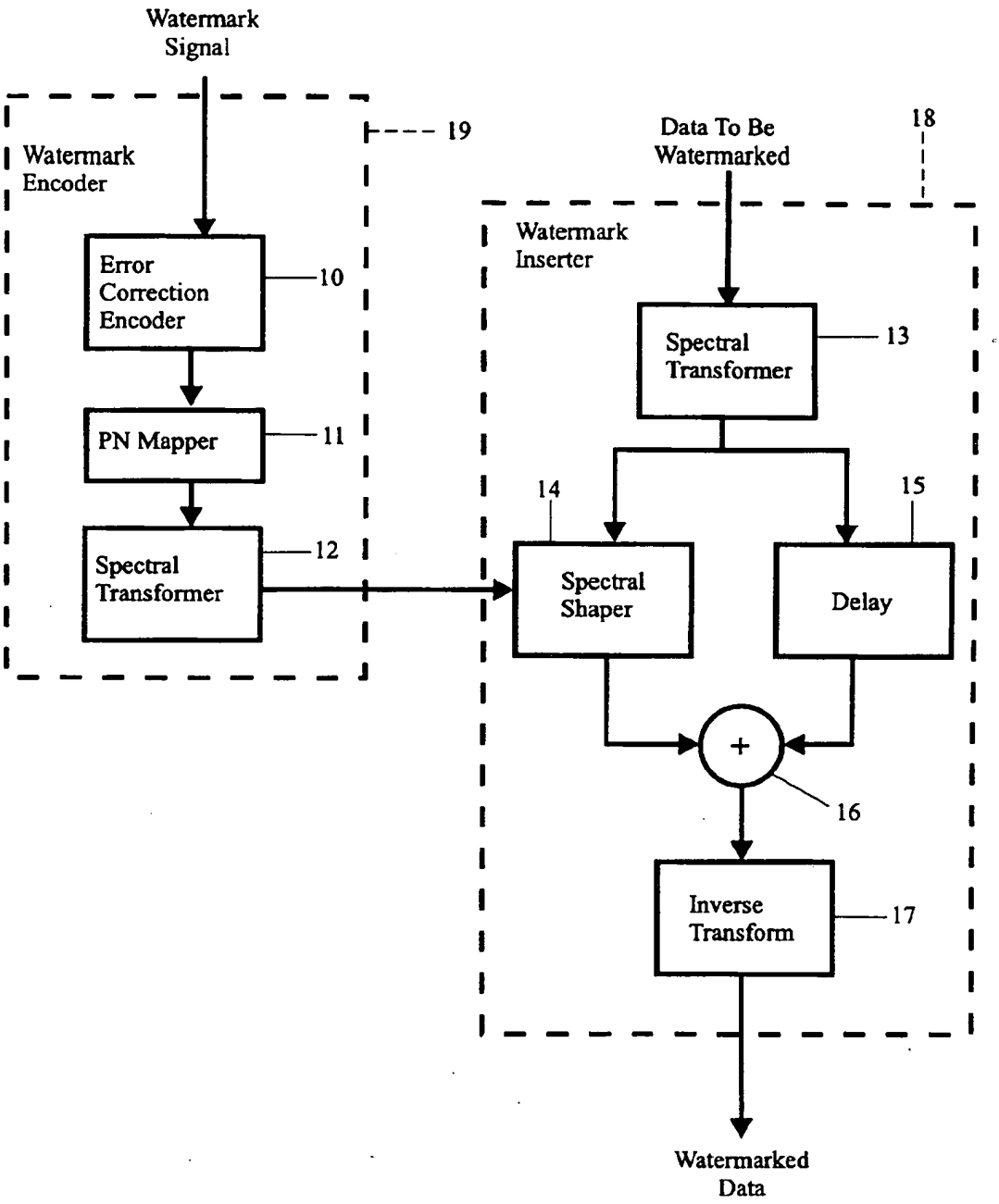

Watermark Signal
Watermark Encoder
19
Error Correction Encoder
10
PN Mapper
11
Spectral Transformer
12
Data To Be Watermarked
18
Watermark Inserter
Spectral Transformer
13
14
15
Spectral Shaper
Delay
+
16
Inverse Transform
17
Watermarked Data

Figure 1

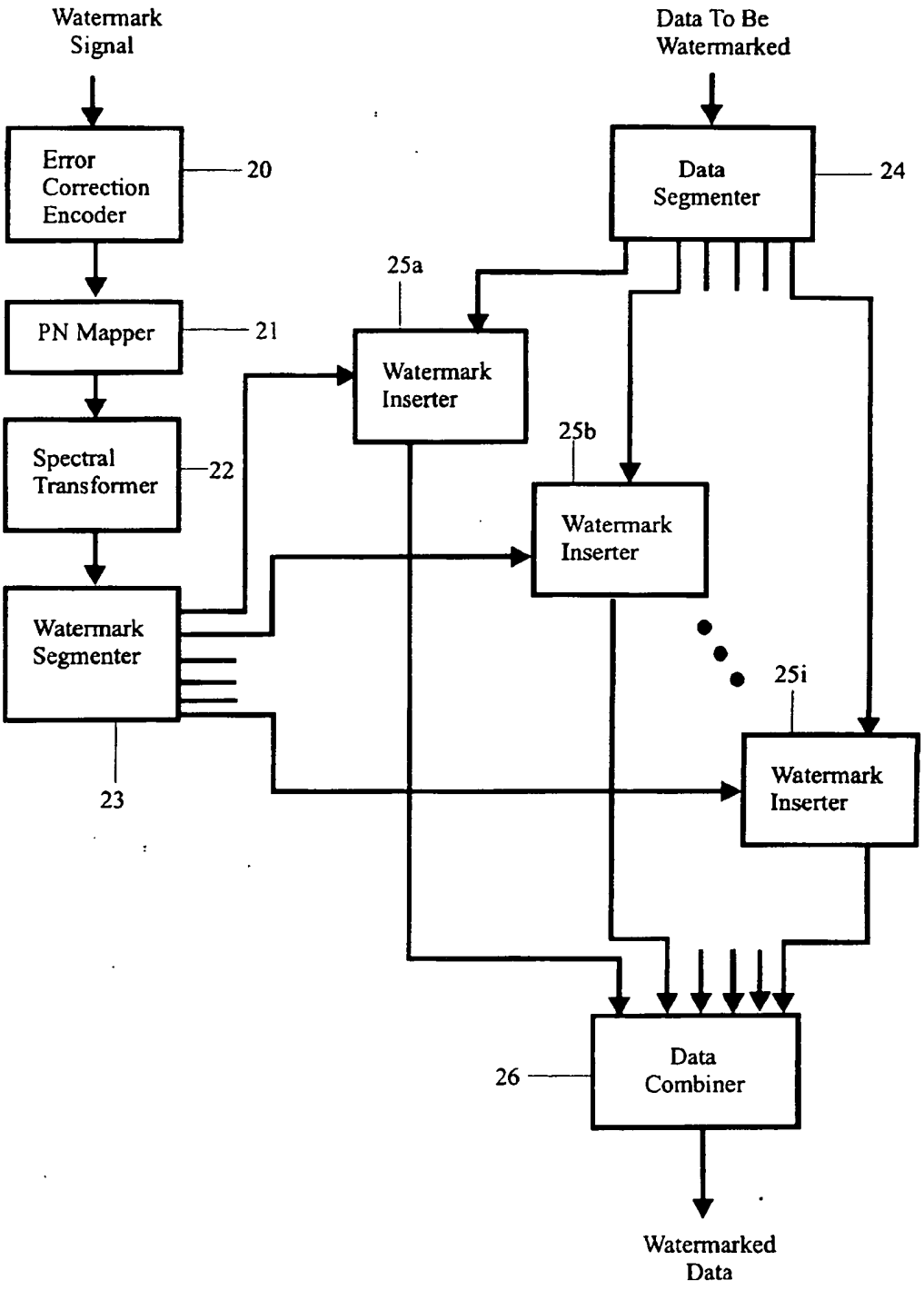

Watermark Signal
Data To Be Watermarked
Error Correction Encoder
20
Data Segmenter
24
25a
PN Mapper
21
Watermark Inserter
25b
Spectral Transformer
22
Watermark Inserter
Watermark Segmenter
25i
Watermark Inserter
23
Data Combiner
26
Watermarked Data

Figure 2

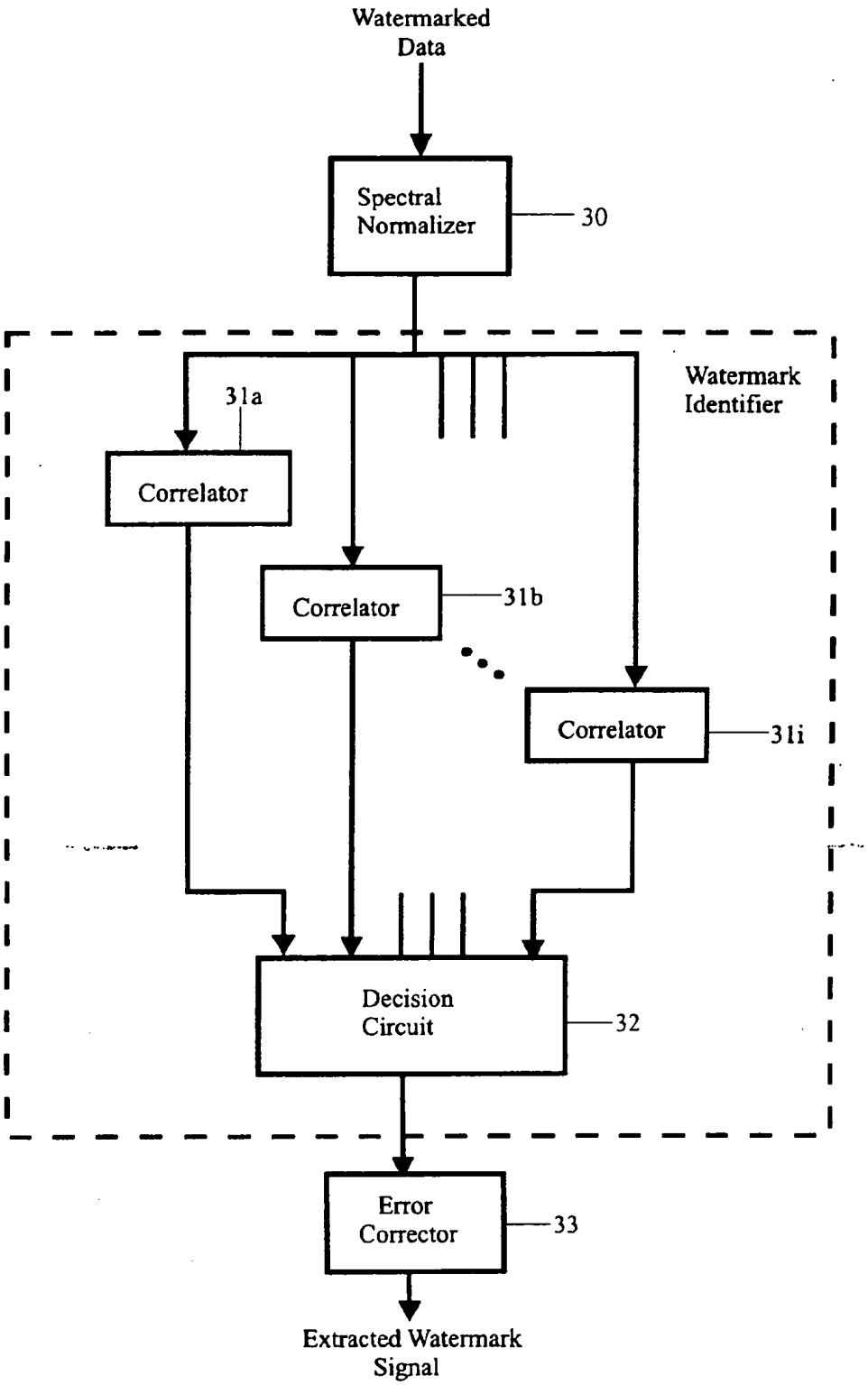

Watermarked Data
Spectral Normalizer
30
Watermark Identifier
31a
Correlator
Correlator
31b
Correlator
31i
Decision Circuit
32
Error Corrector
33
Extracted Watermark Signal

Figure 3

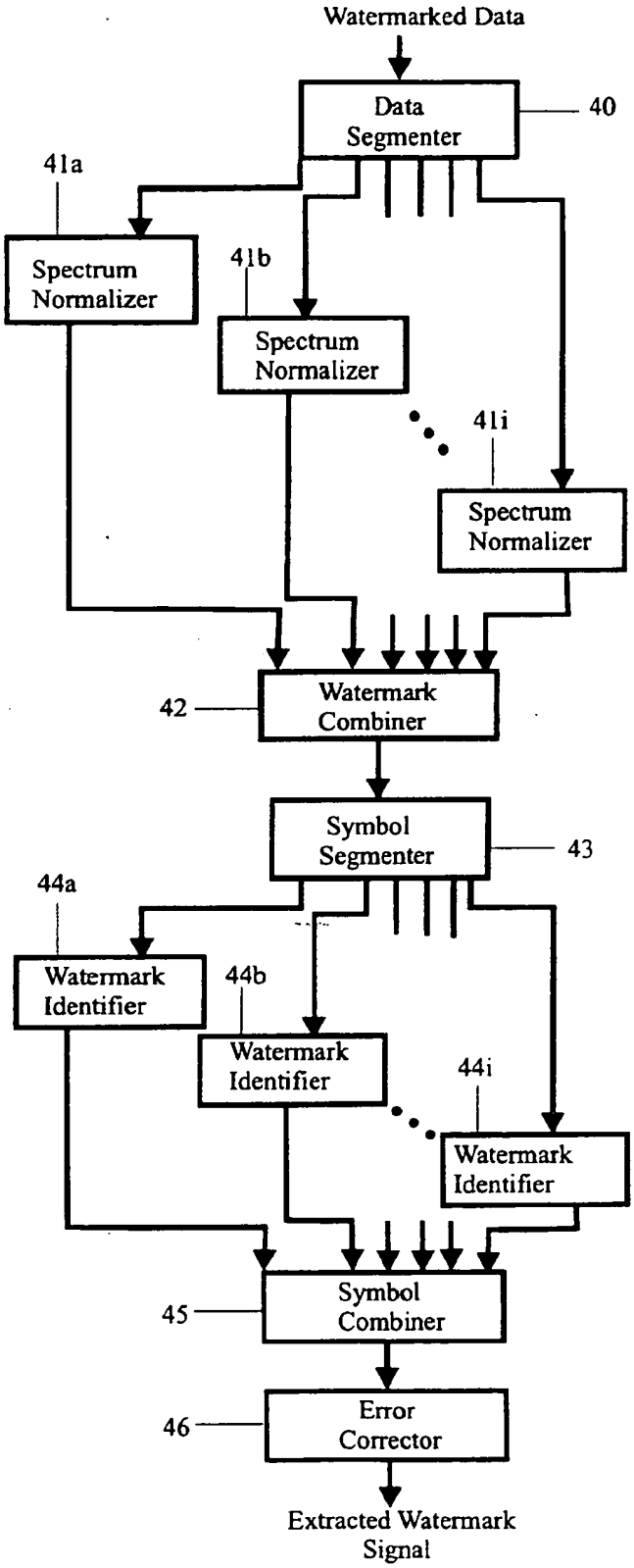

Watermarked Data
Data Segmenter
40
41a
Spectrum Normalizer
41b
Spectrum Normalizer
41i
Spectrum Normalizer
42
Watermark Combiner
Symbol Segmenter
43
44a
Watermark Identifier
44b
Watermark Identifier
44i
Watermark Identifier
45
Symbol Combiner
46
Error Corrector
Extracted Watermark Signal

Figure 4

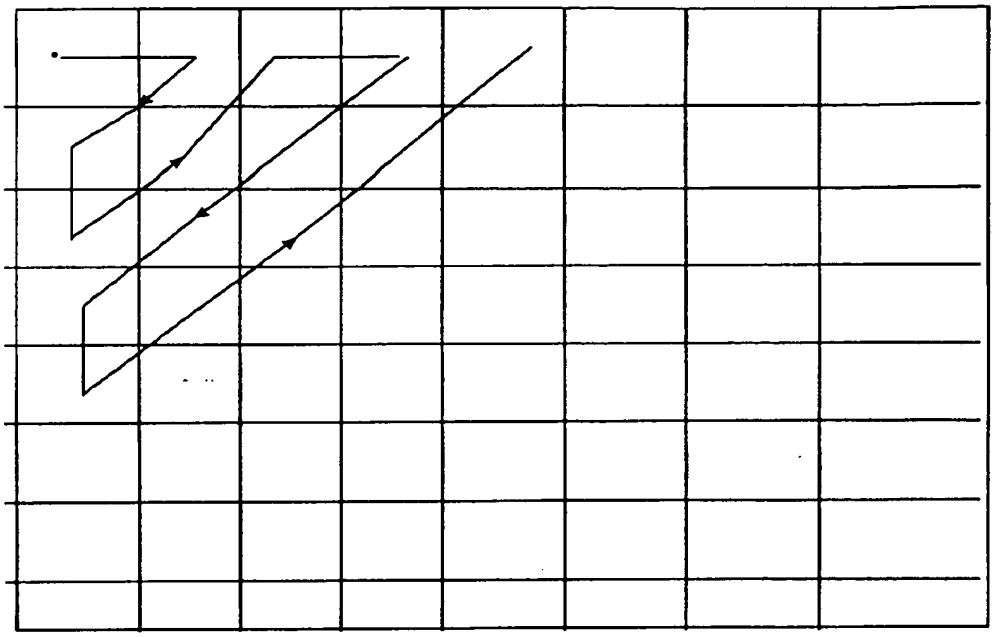


Figure 5

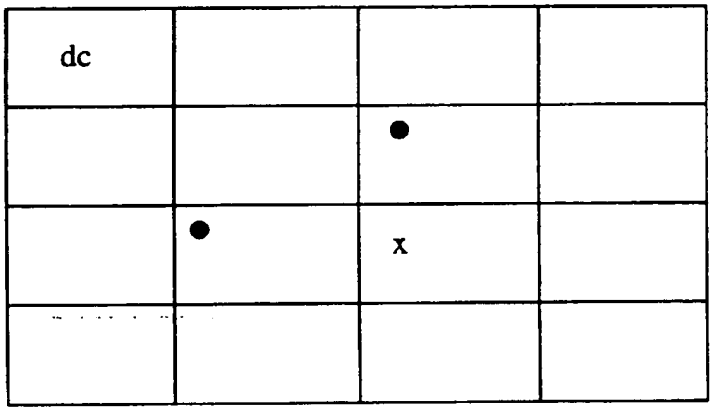


Figure 7

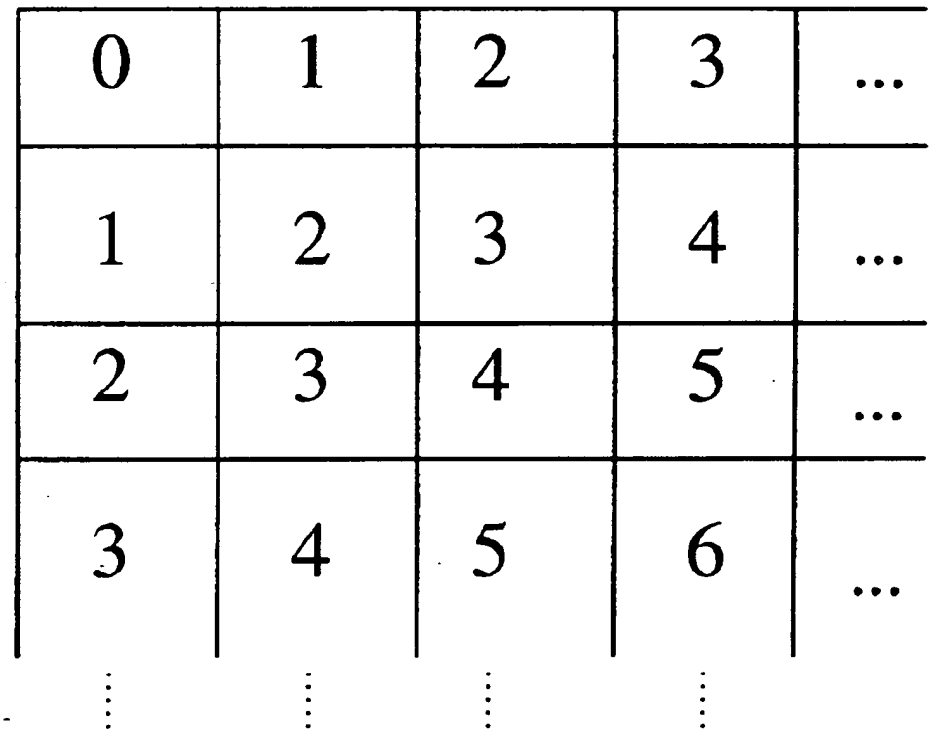


Figure 6

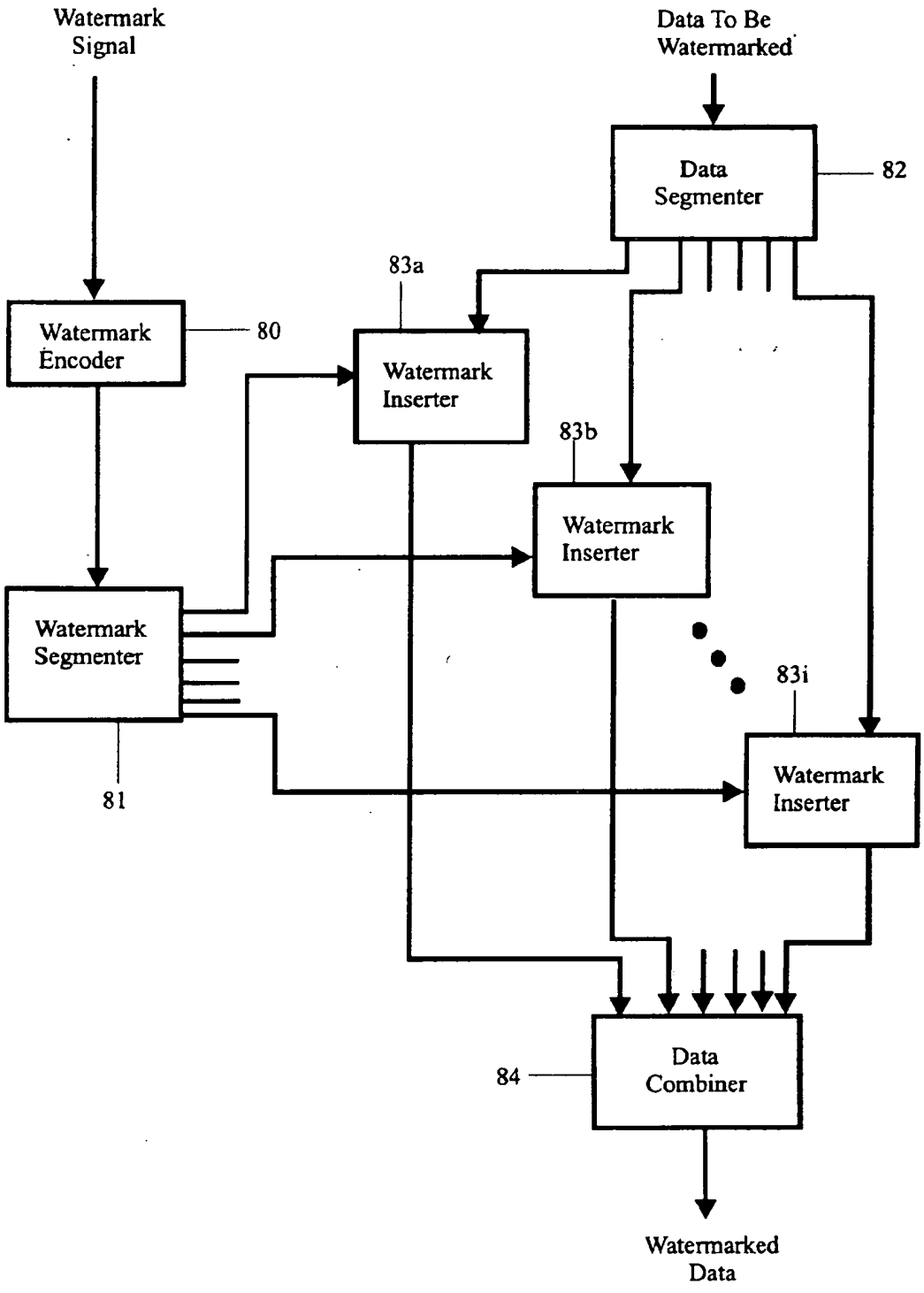

Watermark Signal
Data To Be Watermarked
Data Segmenter
82
83a
Watermark Encoder
80
Watermark Inserter
83b
Watermark Inserter
Watermark Segmenter
83i
Watermark Inserter
81
84
Data Combiner
Watermarked Data

Figure 8

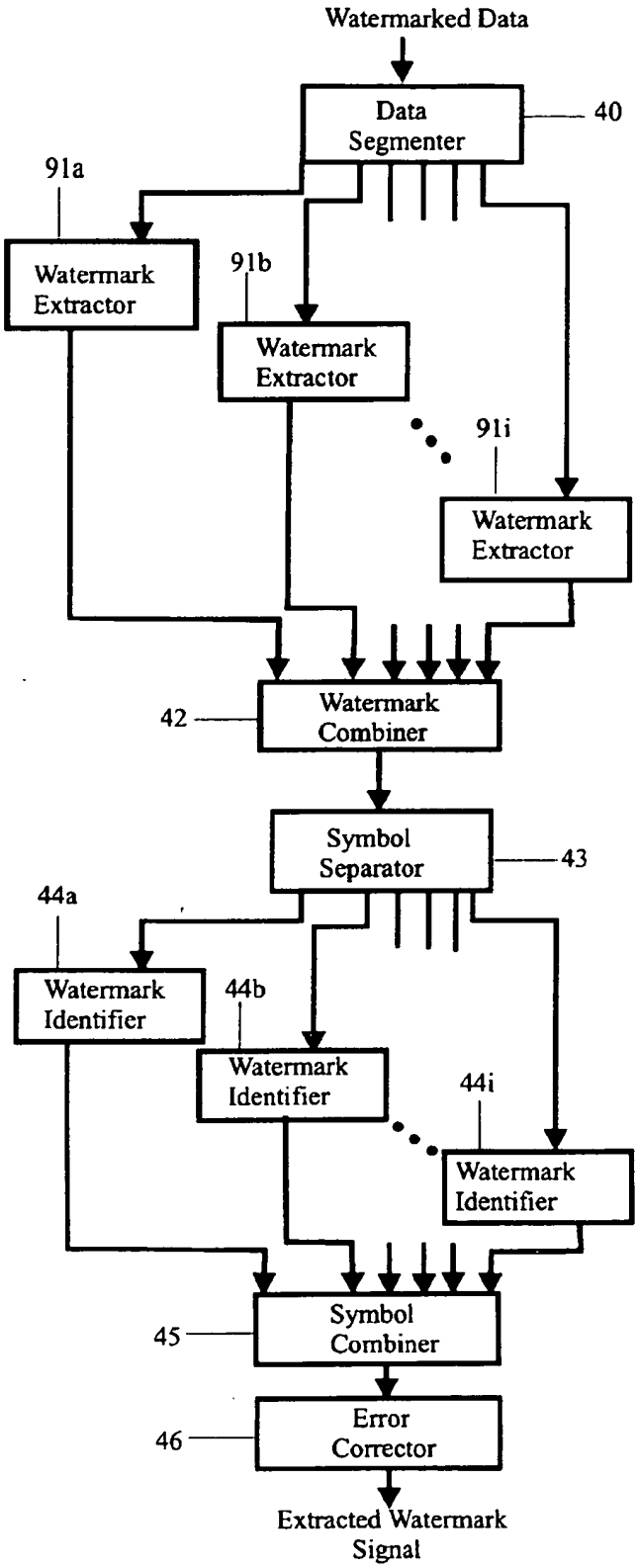

Watermarked Data
Data Segmenter
40
91a
Watermark Extractor
91b
Watermark Extractor
91i
Watermark Extractor
42
Watermark Combiner
Symbol Separator
43
44a
Watermark Identifier
44b
Watermark Identifier
44i
Watermark Identifier
45
Symbol Combiner
46
Error Corrector
Extracted Watermark Signal

Figure 9

DIGITAL WATERMARKING

FIELD OF INVENTION

The present invention relates to digital watermarking of data including image, video and multimedia data. Specifically, the invention relates to the insertion and extraction of embedded signals for purposes of watermarking, in which the insertion and extraction procedures are repeatedly applied to subregions of the data. When these subregions correspond to the 8x8 pixel blocks used for MPEG and JPEG compression and decompression, the watermarking procedure can be tightly coupled with these compression algorithms to achieve very significant savings in computation.

BACKGROUND OF THE INVENTION

The proliferation of digitized media such as image, video and multimedia is creating a need for a security system which facilitates the identification of the source of the material.

Content providers, i.e. owners of works in digital data form, have a need to embed signals into video/image/multimedia data which can subsequently be detected by software and/or hardware devices for purposes of authenticating copyright ownership, control and management.

For example, a coded signal might be inserted in data to indicate that the data should not be copied. The embedded signal should preserve the image fidelity, be robust to common signal transformations and resistant to tampering. In addition, consideration must be given to the data rate that can be provided by the system, though current requirements are relatively low—a few bits per frame.

In U.S. patent application Ser. No. 08/534,894, filed Sep. 28, 1995,. entitled "Secure Spread Spectrum Watermarking for Multimedia Data" now abandoned and assigned to the same assignee as the present invention, which is incorporated herein by reference, there was proposed a spread spectrum watermarking method which embedded a watermark signal into perceptually significant regions of an image for the purposes of identifying the content owner and/or possessor. A strength of this approach is that the watermark is very difficult to remove. In fact, this method only allows the watermark to be read if the original image or data is available for comparison. This is because the original spectrum of the watermark is shaped to that of the image through a non-linear multiplicative procedure and this spectral shaping must be removed prior to detection by matched filtering and the watermark is inserted into the N largest spectral coefficients, the ranking of which is not preserved after watermarking. Thus, this method does not allow software and hardware devices to directly read embedded signals.

In an article by Cox et al., entitled "Secured Spectrum Watermarking for Multimedia" available at http://www.neci.nj.com/tr/index.html (Technical Report No. 95-10) spread spectrum watermarking is described which embeds a pseudo-random noise sequence into the digital data for watermarking purposes.

The above prior art watermark extraction methodology requires the original image spectrum be subtracted from the watermark image spectrum. This restricts the use of the method when there is no original image or original image spectrum available. One application where this presents a significant difficulty is for third party device providers desiring to read embedded information for operation or denying operation of such a device.

In U.S. Pat. No. 5,319,735 by R. D. Preuss et al entitled "Embedded Signalling" digital information is encoded to produce a sequence of code symbols. The sequence of code symbols is embedded in an audio signal by generating a corresponding sequence of spread spectrum code signals representing the sequence of code symbols. The frequency components of the code signal being essentially confined to a preselected signaling band lying within the bandwidth of the audio signal and successive segments of the code signal corresponds to successive code symbols in the sequence. The audio signal is continuously frequency analyzed over a frequency band encompassing the signalling band and the code signal is dynamically filtered as a function of the analysis to provide a modified code signal with frequency component levels which are, at each time instant, essentially a preselected proportion of the levels of the audio signal frequency components in corresponding frequency ranges. The modified code signal and the audio signal are combined to provide a composite audio signal in which the digital information is embedded. This component audio signal is then recorded on a recording medium or is otherwise subjected to a transmission channel. Two key elements of this process are the spectral shaping and spectral equalization that occur at the insertion and extraction stages, respectively, thereby allowing the embedded signal to be extracted without access to the unwatermarked original data.

In U.S. patent application Ser. No. 08/708,331, filed Sep. 4, 1996,entitled "A Spread Spectrum Watermark for Embedded Signaling" by Cox; now U.S. Pat. No. 5,848,155 and incorporated herein by reference, there is described a method for extracting a watermark of embedded data from watermarked images or video without using an original or unwatermarked version of the data. This work can be viewed as an extension of the original work of Preuss et al from the audio domain to images and video.

This method of watermarking an image or image data for embedding signaling requires that the DCT (discrete cosine transform) and its inverse of the entire image be computed. There are fast algorithms for computing the DCT in N log N time, where N is the number of pixels in the image. However, for N=512x512, the computational requirement is still high, particularly if the encoding and extracting processes must occur at video rates, i.e. 30 frames per second. This method requires approximately 30 times the computation needed for MPEG-II decompression.

One possible way to achieve real-time video watermarking is to only watermark every N^{th} frame. However, content owners wish to protect each and every video frame. Moreover, if it is known which frames contain embedded signals, it is simple to remove those frames with no noticeable degradation in the video signal.

In U.S. patent application Ser. No. 08/715,953, filed Sep. 19, 1996,entitled "Watermarking of Image Data Using MPEG/SPEG Coefficients" by Cox, and incorporated herein by reference, there is described an alternative method, which is to insert the watermark into nxn blocks of the image (subimages) where n<<N. Then the computation cost is

$$\frac{N}{n} n \log n = N \log n.$$

For $N=512\times512=2^{18}$ and $n=8\times8=2^6$, the asymptotic saving is only a factor of 3. However, empirically the cost of computing the DCT over the entire image may be significantly higher when cache, loop unfolding and other efficiency issues are considered. Thus, the practical difference

may approach a 30 fold savings. More importantly, if the block size is chosen to be 8x8, i.e. the same size as that used for MPEG image compression, then it is possible to tightly couple the watermark insertion and extraction procedures to those of the MPEG compression and decompression algorithms. Considerable computational saving can then be achieved since the most expenses computations relate to the calculation of the DCT and its inverse and these steps are already computed as part of the compression and decompression algorithm. The incremental cost of watermarking is then very small, typically less than 5% of the computational requirements associated with MPEG.

The present invention improves the reliability of the invention described in the 08/715,953 application, now pending by storing watermark information into subimages, and extracting watermark information from subimages, in a manner different from that described earlier.

SUMMARY OF THE INVENTION

The present invention improves the reliability of the prior systems by systematically varying the order in which watermark signal components are inserted into each subimage, by inserting only part of the watermark signal into each subimage, and, during watermark detection, by combining the watermark signals found in groups of subimages to reconstruct the original watermark signal before testing for correlation with any predefined watermarks.

For detection, a reverse transformation is applied to each subimage to reconstruct the watermark information that was stored in that subimage. The resulting signals are then averaged together to reconstruct the whole watermark, and to reduce noise. Finally, this reconstructed watermark is compared against a predefined set of watermark signals to determine which one was inserted into the image.

A principal object of the present invention is therefore, the provision of inserting a subset of a watermark into a subset of subregions of data to be watermarked.

Another object of the invention is the provision of a digital watermarking system in which a watermark is extracted by averaging the watermarked signal from subregions of watermarked data, and then correlating the resulting signal to determine the watermark.

A further object of the invention is the provision of a digital watermarking system in which the watermark is composed of two portions, a verification portion and a synchronization portion, in order to improve watermark extraction reliability

Further and still other objects of the invention will become more clearly apparent when the following description is read in conjunction with the accompanying drawing.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a schematic block diagram of watermark insertion procedure;

FIG. **2** is a schematic block diagram of a watermark insertion procedure in accordance with the teachings of the present invention;

FIG. **3** is a schematic block diagram of a watermark extraction procedure;

FIG. **4** is a schematic block diagram of a watermark extraction procedure in accordance with the teachings of the present invention;

FIG. **5** is a graphic representation of a zigzag pattern useful for vectorizing subimages;

FIG. **6** is a graphic representation of rotation of PN sequences;

FIG. **7** is a graphical representation of an 8x8 block shown the spatial relation of averaged terms;

FIG. **8** is a schematic block diagram of a method for inserting watermarks in accordance with the present invention; and

FIG. **9** is a schematic block diagram of a method for extracting watermarks in accordance with the present invention.

DETAILED DESCRIPTION

Referring now to the figures, and FIGS. **1** through **4** in particular, there is shown schematic block diagrams of a general method for inserting and detecting watermarks in digital data, for instance images.

In the following description, reference may be made to image data or images. While the invention has applicability to image data and images, it will be understood that the teachings herein and the invention itself are equally applicable to video, image and multimedia data and the term "image" and "image data" will be understood to include these terms where applicable. As used herein, "watermark" will be understood to include embedded data, symbols, images, instructions or any other identifying information.

In the following description, reference is made to procedures described in U.S. patent application Ser. No. 08/534,894 for inserting and extracting or detecting a watermark in images as INSERT-ORIGINAL and EXTRACT-ORIGINAL, respectively. Reference is made to procedures described in U.S. patent application Ser. No. 08/708,331 filed Sep. 4, 1996,now U.S. Pat. No. 5,848,155 for inserting and extracting or detecting watermarks in images as INSERT-WHOLE and EXTRACT-WHOLE, respectively. And reference is made to procedures described in U.S. patent application Ser. No. 08/715,953 for inserting and extracting or detecting watermarks in images as INSERT-MPEG-A and EXTRACT-MPEG-A, respectively.

FIG. **1** shows a schematic block diagram of INSERT-WHOLE procedure for inserting watermarks into images. The watermark signal, in the form of a finite sequence of symbols chosen from an alphabet, is provided as an input to an error correction encoder **10** which transforms this sequence into another sequence that contains redundant information. The output of encoder **10** is provided to a PN-mapper **11**, which maps each symbol of the encoded watermark into a pre-specified pseudo-random noise (PN) code. The output of the PN-mapper **11** is provided to a spectral transformer **12**, which converts the pseudo-random noise sequence into the frequency domain. The conversion preferably is by discrete cosine transform (DCT), however, fast fourier transform, wavelet type decomposition and the like may also be used for frequency conversion. Concurrently, the data to be watermarked is provided to another spectral transformer **13**. The outputs of the two spectral transformers **12** and **13** are then provided as inputs to a spectral shaper **14**, which modifies the spectral properties of the pseudo-random noise codes from spectral transformer **12** to mask the watermark when added to the image data. The spectrally transformed data to be watermarked, from spectral transformer **13**, is also provided as an input to a delay **15**. The output of the spectral shaper **14** is then added to the output of delay **15** at a summer **16**. The summer output is subject to an inverse transform **17**. The result of the inverse transform is watermarked data.

INSERT-MPEG-A differs from INSERT-WHOLE by segmenting the data to be watermarked into multiple blocks,

such as 8×8 pixel subimages or subregions. Each block of data then has the watermark inserted according to the above described method. That is, for each 8×8 subimage or subregion, a pseudo-random number (PN) sequence is inserted into the DCT coefficients after suitable spectral shaping. The procedure is repeated for all such subimages or subregions. The size of the subimage or subregion is preferably 8×8, but it can be of other sizes, such as 2×2, 3×3, 4×4 or 16×16.

FIG. 2 shows a schematic block diagram of a watermark insertion procedure in accordance with teachings of the present invention. The watermark signal is processed into a noise spectrum signal by the error correction encoder **20**, the PN mapper 21, and the spectral transformer **22**, in the same manner as described in conjunction with FIG. 1. However, unlike INSERT-WHOLE or INSERT-MPEG-A, the watermark is then used as an input to a watermark segmenter **23**, which systematically separates the watermark into several subwatermarks. Any portion of the original watermark might appear redundantly in several of the resulting subwatermarks. Concurrently, the data to be watermarked is used as an input to data segmenter **24**, which segments the data into blocks or subregions, such as 8×8 subimages, as in INSERT-MPEG-A. Each of the subwatermarks output by the watermark segmenter **23** is then inserted into a data block by one of the watermark inserters **25***a*, **25***b*, etc. The procedure used by the watermark inserters **25***a*, **25***b*, etc., is the same procedure described connection with watermark inserter **18** in FIG. **1**. That is, each subwatermark is added into a spectrally transformed data block after spectral shaping, and the resulting data is then transformed back into the spatial domain. Finally, the watermarked data blocks from the watermark inserters **25***a*,**25***b*, etc., are assembled by data combiner **26** to produce watermarked data.

FIG. **3** shows a schematic block diagram of the EXTRACT-WHOLE procedure. The watermarked image, video or multimedia data is first used as input into a spectral normalizer **30** to undo any previously performed spectral shaping. If the data contains a watermark, then the output of the spectral normalizer **30** will resemble the spectral transformation of the PN coding of that watermark (the signal that was input to the spectral shaper **14** in FIG. **1**). The output of the spectral normalizer **30** is then used as an input to several correlators **31***a*, **31***b*, etc., which test the watermark with the PN codes used to represent the various symbols that the encoded watermark might contain (i.e. each correlator tests for one PN code that is used to encode a symbol by the PN mapper **11** of FIG. 1). The outputs of the correlators **31***a*, **31***b*, etc., are used as inputs to a decision circuit **32**, which determines the most likely sequence of symbols. Finally, this sequence is corrected by an error corrector **33**, which performs the inverse coding that was performed by the error correction encoder **10** in FIG. **1**. The result is the extracted watermark.

In EXTRACT-MPEG-A, the data from which a watermark is to be extracted is first segmented into several blocks, such as 8×8 subimages, exactly as in INSERT-MPEG-A. The signal from each subimage is then normalized and used as input into a bank of correlators similar to the correlators **31***a*, **31***b*, etc. in FIG. **3**. The output from the correlators is then averaged with the outputs of corresponding correlators from other subimages, and the resulting average correlations are used as inputs into the decision circuit **32** for subsequent processing as described above.

FIG. **4** shows a schematic block diagram of a watermark extraction procedure in accordance with the teachings of the present invention. The watermarked data is first segmented into blocks by data segmenter **40**, which corresponds to the data segmenter **24** used during the insertion procedure in FIG. **2**. Each of the data blocks is provided to a respective spectrum normalizer **41***a*, **41***b*, etc. to produce a signal resembling the subwatermark that was inserted into the respective data block. These inserted subwatermark signals are then used as inputs into a watermark combiner **42**. In the combiner **42**, parts of the watermark that appear redundantly in several subwatermarks are averaged together to reduce noise. The output of the watermark combiner **42** is provided as the input to a symbol separator **43** which divides the watermark into parts, each of which corresponds to one symbol from the encoded watermark signal (the output of error correction encoder **20** in FIG. **2**).

These symbols from separator **43** are provided as inputs to respective watermark identifiers **44***a*, **44***b* etc. each of which includes of a bank of correlators and a decision circuit, as shown in FIG. **3**. The outputs of the watermark identifiers are symbols from the alphabet used in the original encoded watermark signal. The identified symbols are reassembled into a complete encoded watermark by the symbol combiner **45**. Finally, the resulting encoded watermark is decoded by the error corrector **46**.

The insertion and extraction procedures will now be described in more detail. In INSERT-ORIGINAL and EXTRACT-ORIGINAL, the object is to embed a single PN (pseudo random number) sequence into an image when the original image is available at the time of extraction. The information associated with the PN sequence is assumed to be stored in a database together with the original image and the spectral location of the embedded watermark. The locations of the watermarked components has to be recorded because the implementation approximated the N perceptually most significant regions of the watermark by the N largest coefficients. However, this ranking was not invariant to the watermarking process. The N largest coefficients may be different after inserting the watermark than before inserting the watermark.

In order to avoid this problem, the present invention places a watermark in predetermined locations of the spectrum, typically the first N coefficients. However, any predetermined locations could be used, though such locations should belong to the perceptually significant regions of the spectrum if the watermark is to survive common signals transformations such as compression, scaling, etc.

More generally, the information to be embedded is a sequence of m symbols drawn from an alphabet A (e.g. the binary digits or the ASCII symbols). This data is then supplemented with additional symbols for error detection and correction. Each symbol is then spread spectrum modulated, a process that maps each symbol into a unique PN sequence known as a chip. The number of bits per chip is preset - the longer the chip length, the higher the detected signal-to-noise ratio will be, but this is at the expense of signaling bandwidth.

The power spectrum of the PN sequence is white, i.e. flat, and is therefore shaped to match that of the "noise", i.e. the image/video/audio/or multimedia data into which the watermark is to be embedded. It is this spectral shaping that must be modified from the prior methods so that the extraction process no longer requires the original image. To do this, each coefficient of the watermarked spectrum is scaled by the local average of the power in the image spectral coefficient rather than the coefficient itself, i.e.

$$f_i' = f_i + \alpha \text{avg}(|f_i|) W_i \quad (1)$$

The averaging is the averaging of the absolute coefficient values and not the coefficient values themselves. This is effectively estimating the average power present at each frequency. Other averaging procedures are possible, for example, averaging over several frames or average of local neighborhoods of 8×8 blocks.

This average may be obtained in several ways. It may be a local average over a two dimensional region. Alternatively, the two dimensional spectrum may be sampled to form a one dimensional vector and a one dimensional local average may be performed. One dimensional vectorization of the two dimensional 8×8 DCT coefficients is already performed as part of MPEG II. The average may be a simple box or weighted average over the neighborhood.

For video data, temporal averaging of the spectral coefficients over several frames can also be applied. However, since several frames are needed for averaging at the spectral normalization stage of the extractor, the protection of individual video frames taken in isolation may not be possible. For this reason, the present invention treats video as a very large collection of still images. In this way, even individual video frames are copy protected.

In order to extract the watermark, it is necessary to perform the spectral normalization, in which the previously performed spectral shaping procedure is inverted. In the present invention, the original unwatermarked signal is not available. Thus, the average power of the frequency coefficients, $\text{avg}(|f_i|)$, is approximated by the average of the watermarked signal, i.e. $\text{avg}(|f_i'|)$

$$\text{avg}(|f_i|) \approx \text{avg}(|f_i'|) \tag{2}$$

This is approximately true since $\alpha\text{avg}(|f_i|)W_i << f_i$, where W_i is the watermark component, and α is a constant typically in the range between 0.1 and 0.01.

The normalization stage then divides each coefficient (f_i') in the received signal by the local average $\text{avg}(|f_i'|)$ in the neighborhood.

That is,

$$\frac{f_i'}{\text{avg}(|f_i'|)} = \frac{f_i + \alpha\text{avg}(|f_i|)W_i}{\text{avg}(|f_i|)} \approx \frac{f_i'}{\text{avg}(|f_i'|)} + \alpha W_i \tag{3}$$

The first term, on the right hand side (RHS) of Equation (3),

$$\frac{f_i}{\text{avg}(|f_i'|)},$$

is considered a noise term. This term was not present in the system described in U.S. patent application Ser. No. 08/534,894, because access to the unwatermarked coefficients allowed this term to be removed. The second term αW_i is the original watermark signal which can now be detected using conventional correlation.

If the watermark is extracted from any single 8×8 block, the detector reliability is very low. If, however, the watermarks extracted from each 8×8 block are first added together and the averaged watermark is then applied to the correlator, then a very strong and unambiguous response is obtained. This differs from the method described in U.S. patent application Ser. No. 08/715,953 in which correlation occurred within each block and the output from each correlator was averaged together. The present invention was found to improve the detection response and significantly reduced the computation requirement associated with each block.

In practicing the present invention preferably there is a unique PN sequence for each symbol in the alphabet. The method is relatively robust to clipping since the detector output reduces linearly with the quantity of 8×8 subimage blocks in the image. For DVD (digital video disk) embedded signaling for APS (analog protection system) and CGMS (copy generation management system), there would be a total of 8 or 16 PN sequences.

The number of 8×8 blocks in a 512×512 image is 4096, suggesting that significantly more than one of 16 symbols can be embedded in an image or video frame. Assume, for example, that it is desired to embed 1 out of 128 symbols in an image. It is necessary to perform 128 parallel correlations. This is computationally tractable but hardware implementations of each correlation become more complex. An alternative method is to only use two binary symbols. It may be preferable to associate more than one PN sequence with each of the two binary symbols or bits in order to increase the difficulty of intentionally removing the watermark. In this case, there are only two correlators and a binary string may be embedded into the image. The raw bit error rate will be very high due by the low detector output. However, this can be reduced to acceptable levels by using error correcting codes, such as Reed-Solomon (RS). RS codes are robust to burst error which may occur because of clipping of the image. Other error correcting codes may also be used.

When using this method, it is necessary for the receiver to know the start location of the encoded block. The start location may not be obvious, particularly when the image has been subjected to clipping. However, convention synchronizing methods can be used; such as preceding each block with a special or unique symbol or string of symbols.

To insert a watermark, each 8×8 block is treated as an individual subimage or subregion. The DCT of the subimage is then computed and the two dimensional DCT is vectorized in the zigzag pattern shown in FIG. 5, although other patterns are also possible. These two stages constitute most of the calculations but are part of the MPEG encoding process. Next, a PN noise sequence $\{w_1 \ldots w_n\}$ is inserted into the DCT coefficients using Equation 1 as before. The length of the PN sequence cannot exceed 64 (in an 8×8 block) and is typically much shorter, in the range of 11 to 25. If only a single code is to be inserted into the image, then the same PN sequence is inserted into each of the 720 × 480/64 −5400 blocks. However, a variation may be performed at this point in the procedure. Within each row of blocks, the PN sequence is cyclically rotated by one frequency coefficient prior to insertion in the subsequent block. Similarly, the PN sequence is cyclically rotated by one frequency coefficient at the start of each new row. FIG. 6 illustrates an order of rotations.

The purpose of these rotations or shifts is to improve the response of the watermark extraction stage. Earlier experiments revealed that certain DCT coefficients were more difficult to estimate than others. The location at these coefficients varied from image to image. However, within an image, the coefficient could be consistently poor. Consequently, without shifting, one or more of the estimated watermark coefficients could be significantly degraded relative to the other watermark coefficients, thereby reducing the detector performance. Conversely, shifting significantly reduces the effect a poor DCT coefficients has on a single watermark coefficient and the detector performance is markedly improved. Note that any cyclic pattern can be used.

Further modifications are useful once rotation of the watermark has been introduced. First, the length of the watermark may now be significantly greater than 64.Then, for each block only a small subset of the watermark (say five) coefficients is inserted into the first five DCT coefficients (excluding the d.c. term). Because of the rotation, a different subset of the watermark is inserted into neighboring 8×8 blocks. Finally, having completed the watermark insertion, the MPEG encoder is able to proceed with the subsequent stages of compression.

Note that the watermark may also be inserted after the MPEG quantization stage to reduce distortion of the watermark. MPEG-2 performs a convenient one dimension vectorization called "zigzagging", which allows a simple 3×1 box average to be performed on the coefficients (excluding the d.c. term).

In practice, performance was improved if the averaging is performed using the 2 four-connected coefficients closest to the d.c. term, as illustrated in FIG. 7, i.e. the two coefficients above and to the left.

Watermark detection begins by first extracting the PN noise sequence from each 8×8 block using Equation 1.For each block, the PN sequence is then cyclically shifted in the opposite direction by one frequency coefficient, and the average over all the blocks is then computed. In practice, this process can be computed incrementally and does not require temporary storage of all the extracted watermarks. A weighted averaging can also be applied, where the weights are determined based on their susceptibility to common signal transformations such as low pass filtering. Finally, the average watermark is compared with the original PN sequence via correlation. The reason for shifting the watermark in the column direction may now be apparent. If the image is clipped on an arbitrary block boundary, then the computed average watermark will simply be rotated by an amount that is a function of the relative location of the clipped portion of the image. Correlation can then be performed on all permutations (typically 11 to 25) of the watermark. The output from the correlator with the maximum value is then used for decision purposes. The extraction stage is depicted in FIG. 4. Taking the maximum correlator output over all rotations of the watermark can cause the decision circuitry to be noisy. To improve this, the watermark is broken into two pieces; a synchronization portion is of length K and a verification portion is N-K. Then, when the watermark is extracted as before, correlation is first performed only on all rotations of the synchronization portion of this watermark. The maximum correlation output is noted, then the verification portion of the watermark is rotated by the corresponding amount and a second correlation is performed on the verification portions of the watermarks. This process significantly improves the overall reliability of the system. In the course of experimentation, it was noticed that some watermarks performed better than others on the same imagery. This was caused by variation in the correlation statistics between the synchronization and verification portions of the watermark. Ideally, the two portions should have very low correlations. However, in several cases where watermarks performed poorly, it was traced to unexpected correlations between the two portions.

The present invention provides a modification to digital watermarking methods in which the original data is required for watermark extraction thereby enabling watermarking extraction in the absence of an unwatermarked or original data. The present invention preferably uses MPEG/JPEG coefficients. An image is divided into typically 8×8 block subimages or subregions and each subimage is processed and the results are combined to derive the extracted watermark. The result is extraction of the watermark with very high confidence.

While the above invention describes improvements to the prior-art INSERT-WHOLE, INSERT-MPEG-A, EXTRACT-WHOLE, and EXTRACT-MPEG-A algorithms, it should be apparent to anyone skilled in the art that the same improvements may be applied to any algorithm for inserting and extracting watermarks in image data. This more general view of the present invention is shown in FIGS. **8** and **9**.

FIG. **8** shows a schematic block diagram of the general method for inserting watermarks. This general method makes use of a non-block-based watermark insertion algorithm, which shall be referred to hereafter as the "base insertion algorithm". The watermark encoder **80** converts the watermark into a form appropriate for the base insertion algorithm. If the base insertion algorithm is that shown in FIG. **1**, for example, then the watermark encoder **80** corresponds to the watermark encoder **19**, which comprises the error correction encoder **10**, the PN mapper **11**, and the spectral transformer **12**. However, if a different base insertion algorithm is to be used, then the watermark encoder **80** may perform a different transformation of the watermark. The encoded watermark signal from watermark encoder **80** is provided as an input to watermark segmenter **81**, which divides the watermark into a set of subwatermarks. Any portion of the original watermark might appear redundantly in several of the resulting subwatermarks. The data to be watermarked is provided as an input to data segmenter **82**, which divides the data into subregions. Each subwatermark is inserted into a respective data subregion by a watermark inserter **83***a*, **83***b*, etc. The watermark inserters implement the base insertion algorithm, so, if the base insertion algorithm is that shown in FIG. **1**, then each watermark inserter **83***a*, **83***b*, etc., corresponds to the watermark inserter **18**, which comprises a spectral transformer **13**, a spectral shaper **14**, a delay **15**, a summer **16**, and an inverse transform **17**. However, if a different base insertion algorithm is to be used, then the watermark inserters **83***a*, **83***b*, etc., may employ a different method of inserting subwatermarks into the subregions of the data to be watermarked. The outputs from the watermark inserters are assembled in data combiner **84** to provide watermarked data.

FIG. **9** shows a schematic block diagram of the corresponding general extraction algorithm. The algorithm makes use of a "base extraction" algorithm that corresponds to the base insertion algorithm used in inserting the watermark into the data to be watermarked (FIG. **8**). The algorithm in FIG. **9** is substantially the same as the algorithm shown in FIG. **4**, except that, in the general case, the spectrum normalizers **41***a*, etc. are replaced by watermark extractors **91***a*, etc., which implement the base extraction algorithm. That is, if the base insertion algorithm used was the algorithm shown in FIG. **1**, then the watermark extractors **91***a*, etc., in FIG. **9** will be the spectrum normalizers **41***a*, etc. in FIG. **4**.

While there has been described and illustrated a system for inserting a watermark into and extracting a watermark from watermarked data without using an unwatermarked version of the data, it will be apparent to those skilled in the art that variations and modifications are possible without deviating from the broad principles and teachings of the present invention which shall be limited solely by the scope of the claims appended hereto.

What is claimed is:

1. A method for inserting a watermark signal into data to be watermarked comprising the steps of:

dividing data to be watermarked into a plurality of subregions;

computing frequency coefficients of the data to be watermarked in each subregion;

spread spectrum modulating a watermark signal to be inserted by mapping the watermark signal into a PN (pseudo-random noise) sequence;

spectral shaping the PN sequence as a function of the average power in each frequency coefficient of the data; and

inserting each spectral shaped PN sequence into predetermined coefficients in the data in each subregion.

2. A method for inserting a watermark signal into data to be watermarked as set forth in claim 1, where said inserting is performed after the data undergoes MPEG quantization processing.

3. A method for inserting a watermark signal into data to be watermarked as set forth in claim 1, where said frequency coefficients are DCT (discrete cosine transform) coefficients.

4. A method for inserting a watermark signal into data to be watermarked as set forth in claim 3, where each subregion is a 8×8 block of pixels.

5. A method for inserting a watermark signal into data to be watermarked as set forth in claim 4, where said inserting is performed after the data undergoes MPEG quantization processing.

6. A method for inserting a watermark signal into data to be watermarked as set forth in claim 1, where each subregion is a 8×8 block of pixels.

7. A method for inserting a watermark signal into data to be watermarked as set forth in claim 6, where said inserting is performed after the data undergoes MPEG quantization processing.

8. A method for inserting a watermark signal into data to be watermarked as set forth in claim 6, where the frequency coefficients of the watermark signal are rotated prior to inserting of each spectral shaped PN sequence into the subregion.

9. A method for inserting a watermark signal into data to be watermarked as set forth in claim 8, where said inserting is performed after the data undergoes MPEG quantization processing.

10. A method for inserting a watermark signal into data to be watermarked as set forth in claim 8, where only a subset of the watermark signal frequency coefficients is inserted into any one subregion.

11. A method for inserting a watermark signal into data to be watermarked as set forth in claim 10, where the watermark signal comprises a synchronization portion and a verification portion.

12. A method for inserting a watermark signal into data to be watermarked as set forth in claim 11, where said inserting is performed after the data undergoes MPEG quantization processing.

13. A method for inserting a watermark signal into data to be watermarked as set forth in claim 11, where the synchronization portion and the verification portion have very little correlation between each other.

14. A method for inserting a watermark signal into data to be watermarked as set forth in claim 1, where the spectral shaping as a function of the average power is typically 3×1 window of the coefficient obtained from the one-dimensional vectorization by zigzagging of two-dimension frequency coefficients.

15. A method for inserting a watermark signal into data to be watermarked as set forth in claim 1, where the spectral shaping is a function of the average power based on the two four-connected frequency coefficients closest to the DC term.

16. A method of extracting a watermark from watermarked data comprising the steps of:

receiving subregions of watermarked data;

spectrum normalizing the watermarked data as a function of the average power in each frequency coefficient of the watermarked data in each subregion to generate respective normalized signals;

combining the respective normalized signals from each subregion to generate a single watermark;

correlating the single watermark with predetermined PN (pseudo-random noise) sequences corresponding to predetermined symbols to provide correlated signals for each predetermined PN sequence in each subregion;

deciding which correlated signal is most likely a current symbol; and

extracting a sequence of most likely current symbols corresponding to the watermark.

17. A method of extracting a watermark from watermarked data as set forth in claim 16, where the subregions are 8×8 blocks used for MPEG encoding and decoding.

18. A method of extracting a watermark from watermarked data as set forth in claim 17, where said combining the normalized signals from each subregion to generate a single watermark, including removing the relative rotation of the watermark between blocks.

19. A method of extracting a watermark from watermarked data as set forth in claim 18, further comprising subsequently reconstructing the watermark from partial watermarks inserted into each block.

20. A method of extracting a watermark from watermarked data as set forth in claim 19, further comprising weighting the watermark coefficients based on their location within the frequency spectrum, where the weighting is a function of the susceptibility of each frequency coefficient to common signal transformations.

21. A method of extracting a watermark from watermarked data as set forth in claim 16, further comprising correlating with all rotational shifts of the extracted watermark and selecting the maximum value.

22. A method of extracting a watermark from watermarked data as set forth in claim 16, further comprising correlating with all rotational shifts of a synchronization portion of a watermark to determine a maximum value and subsequently rotating a verification portion of the watermark by the same amount as the synchronization portion is rotated to obtain the maximum value prior to correlating between the verification portion and predetermined PN sequences.

23. A method of extracting a watermark from watermarked data comprising the steps of:

receiving subregions of watermarked data;

spectrum normalizing the watermarked data as a function of the average power in each frequency coefficient of the watermarked data in each subregion to generate respective normalized signals;

correlating the respective normalized signals with predetermined PN sequences corresponding to predetermined symbols to provide correlated signals for each predetermined PN sequence in each subregion;

deciding which correlated signal is most likely a current symbol in each subregion for providing an extracted symbol stream;

error correcting the extracted symbol stream; and

extracting a sequence of most likely current symbols corresponding to the watermark.

24. A method of extracting a watermark from watermarked data as set forth in claim 23, where said error correction is Reed Solomon error correction.

25. A method for inserting a watermark signal into data to be watermarked comprising the steps of:

dividing data to be watermarked into a plurality of subregions;

dividing a watermark signal into a plurality of subwatermarks where portions of the watermark are contained in more than one subwatermark; and

inserting said plurality of subwatermarks into said plurality of subregions.

26. A method for inserting a watermark signal into data to be watermarked as set forth in claim 25, where each subwatermark is inserted into a respective subregion, so that each subregion contains at least one subwatermark.

27. A method for extracting a watermark signal from watermarked data comprising the steps of:

receiving a plurality of subregions of watermark data;

extracting a subwatermark from each subregion of said plurality of subregions; and

combining and averaging the subwatermarks extracted from all the subregions to obtain a signal commensurate with the watermark signal.

28. A method for extracting a watermark signal from watermarked data as set forth in claim 27, further comprising the steps of:

dividing the signal commensurate with the watermark signal into a plurality of symbol signals;

correlating each symbol signal with a set of predefined signals;

determining which predefined signal best corresponds to each symbol signal; and

combining the best corresponding predetermined signals to generate the watermark signal.

* * * * *

US005930369A

United States Patent [19]

Cox et al.

[11] Patent Number: 5,930,369

[45] Date of Patent: Jul. 27, 1999

[54] SECURE SPREAD SPECTRUM WATERMARKING FOR MULTIMEDIA DATA

[75] Inventors: Ingemar J. Cox, Lawrenceville; Joseph J. Kilian, Princeton Junction; Talal G. Shamoon, Princeton, all of N.J.

[73] Assignee: NEC Research Institute, Inc., Princeton, N.J.

[21] Appl. No.: 08/926,720

[22] Filed: Sep. 10, 1997

Related U.S. Application Data

[63] Continuation of application No. 08/534,894, Sep. 28, 1995, abandoned.

[51] Int. Cl.[6] G09C 5/00; H04L 9/00

[52] U.S. Cl. 380/54; 380/3; 380/4; 380/23; 380/55; 283/73; 283/113; 283/17

[58] Field of Search 380/3, 4, 9, 23, 380/54, 59, 51, 55; 283/73, 113, 17, 901

[56] References Cited

U.S. PATENT DOCUMENTS

4,939,515	7/1990	Adelson	341/51
5,319,735	6/1994	Preuss et al.	395/2.14
5,488,664	1/1996	Shamir	380/54
5,530,751	6/1996	Morris	380/4
5,530,759	6/1996	Braudaway	380/54
5,568,570	10/1996	Rabbani	382/238
5,646,997	7/1997	Barton	380/23
5,659,726	8/1997	Sandford, II et al.	395/612
5,664,018	9/1997	Leighton	380/54
5,734,752	3/1998	Knox	380/54 X

FOREIGN PATENT DOCUMENTS

0690595	1/1995	European Pat. Off. .	
2196167	4/1998	United Kingdom .	
8908915	9/1989	WIPO	G11B 20/10
9514289	5/1995	WIPO	G06K 19/14
9520291	7/1995	WIPO .	
9621290	7/1996	WIPO	H04H 1/00
9625005	8/1996	WIPO	H04H 7/08
9627259	9/1996	WIPO .	

OTHER PUBLICATIONS

I. Cox et al, "Secure Spread Spectrum Watermarking for Images, Audio and Video", in IEEE Int. Conference On Image Processing, vol. 3, pp. 243–246, 1996.

I. Cox et al, "A Secure, Robust Watermark for Multimedia", in Information Hiding: First Int. Workshop Proc., R. Anderson, ed., vol. 1174 of Lecture notes in Computer Science, pp. 185–206, Springer–Verlag 1996 IEEE Int. Conf. On Image Processing, 1996.

J. Brassil et al, "Watermarking document images with bounding box expansion", in Information Hiding, R. Anderson, ed., vol. 1174, of Lecture Notes in Computer Science, pp. 227–235, Springer–Verlag, 1996.

J.R. Smith et al, "Modulation and information hiding in images", in information Hiding: First Int. Workshop Proc., R. Anderson, ed., vol. 1174 of Lecture Notes in Computer science, pp. 207–226, Springer–Verlag 1996.

R.L. Rivest et al, "A method for obtaining digital signatures and public–key cryptosystems", Communications Of the ACM, vol. 21, No. 2, Feb. 1978, pp. 120–126.

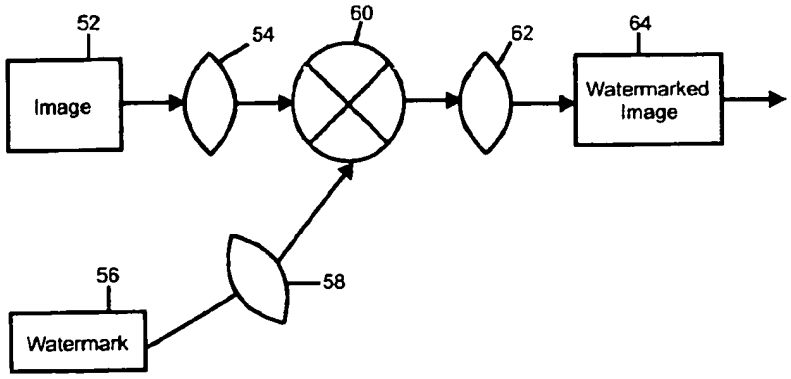
(List continued on next page.)

Primary Examiner—Bernarr E. Gregory
Attorney, Agent, or Firm—Philip J. Feig

[57] ABSTRACT

Digital watermarking of audio, image, video or multimedia data is achieved by inserting the watermark into the perceptually significant components of a decomposition of the data in a manner so as to be visually imperceptible. In a preferred method, a frequency spectral image of the data, preferably a Fourier transform of the data, is obtained. A watermark is inserted into perceptually significant components of the frequency spectral image. The resultant watermarked spectral image is subjected to an inverse transform to produce watermarked data. The watermark is extracted from watermarked data by first comparing the watermarked data with the original data to obtain an extracted watermark. Then, the original watermark, original data and the extracted watermark are compared to generate a watermark which is analyzed for authenticity of the watermark.

46 Claims, 6 Drawing Sheets



OTHER PUBLICATIONS

E. Franz et al, "Computer-based steganography: how it works and why therefore any restrictions on Crytography are nonsense, at best" in Information Hiding, First Int. workshop, Cambridge, UK. Jun. 1996.

Ingemar J. Cox et al., "A Review of watermarking and the importance of perceptual modeling" Proc. of IE '97, vol. 3016, Feb. 9–14, 1997.

Boland et al., "Watermarking Digital Images for Copyright Protection", Image Processing and Its Applications, Jul. 4–6 1995, Conference Publication No. 410.

Kahn et al., "Information Hiding—An Annotated Bibliography",Oct. 17, 1995.

R.G. Van Schyndel et al, "A digital watermark," in Intl. Conf. On Image Processing, vol. 2, pp. 86–90, 1994.

G. Caronni, "Assuring Ownership Rights for Digital Images," in Proc. Reliable IT Systems, VIS '95, 1995.

J. Brassil et al, "Electronic Marking and Identification Techniques to Discourage Document Copying," in Proc. Infocom '94, pp. 1278–1287, 1994.

K. Tanaka et al, "Embedding Secret Information into a Dithered Multi-Level Image," in IEEE Military Comm. Conf., pp. 216–220, 1990.

K. Mitsiu et al, "Video–Steganography: How to Secretly Embeded a Signature in a Picture," in IMA Intelluctual Property Project Proc., vol. 1, pp. 187–206, 1994.

Macq and Quisquater, "Cryptology for Digital TV Broadcasting," in Proc. of the IEEE, vol. 83, No. 6, pp. 944–957, 1995.

W. Bender et al, "Techniques for data hiding," in Proc. of SPIE, vol. 2420, No. 40. Jul. 1995.

Koch, Rindfery and Zhao, "Copyright Protection for Multimedia Data," in Proc. of the Int'l Conf. on Digital Media and Electronic Publishing (Leeds, UK, Dec. 6–8, 1994).

Kock and Zhao, "Towards Robust and Hidden Image Copyright Labeling," in Proc. of 1995 IEEE Workshop on Nonlinear Signal and Image Processing (Neos Marmaras, Halkidiki, Greece, Jun. 20–22, 1995.

Zhao and Koch, " Embedding Robust Labels Into Images For Copyright Protection,"in Proc. Int. Congr. on IPR for Specialized Information, Knowledge and New Technologies (Vienna, Austria), Aug. 21–25, 1995.

"Digital Copyright: Who Owns What?" NewMedia, Sep. 1995, pp. 38–43.

"Publish and Be Robbed?" New Scientist, Feb 18, 1995, pp. 32–37.

Kohno et al, "Spread Spectrum Access Methods for Wireless IEEE Communications," in Communications Magazine, Jan. 1995, pp. 58–67, 116.

Campana and Quinn, "Spread spectrum communications," in IEEE Potentials. Apr. 1993, pp. 13–16.

Mowbray and Grant, "Wideband coding for uncoordinated multiple access communications," in Electronics & Communication Engineering Journal, Dec. 1992, pp. 351–361.

Digimarc Overview & "Wired" Magazine article (Jul. 1995 issue)—Jun. 1995.

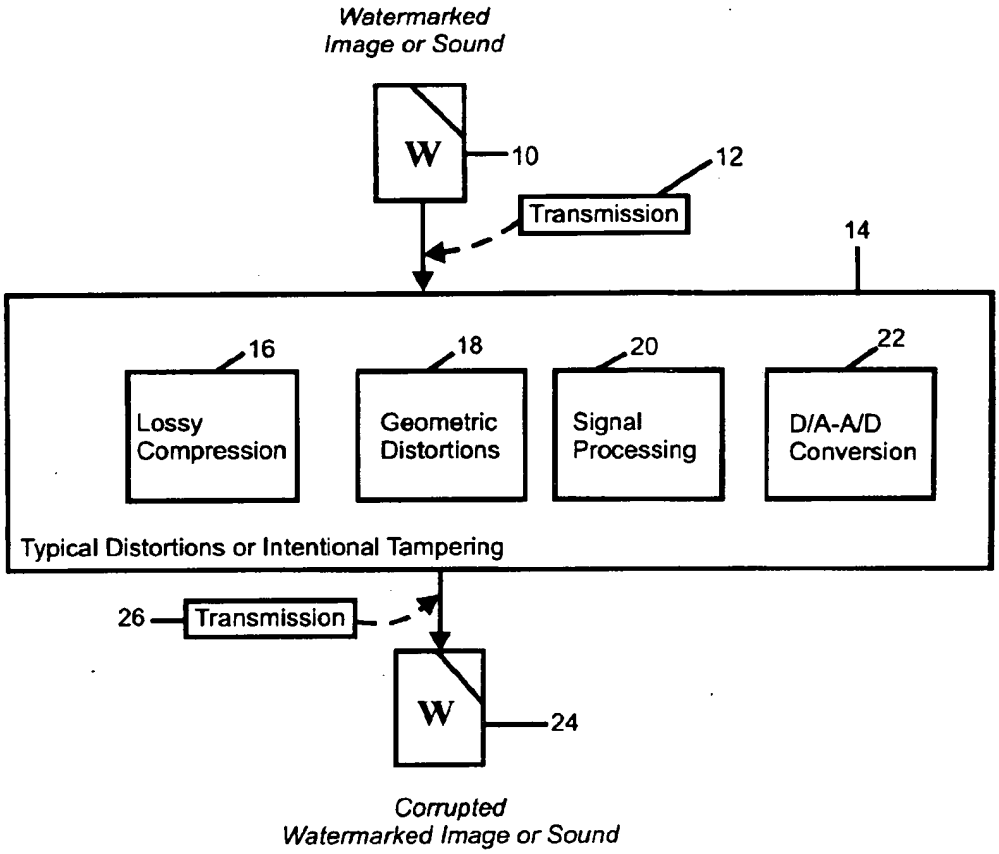

Watermarked
Image or Sound
W
10
12
Transmission
14
16
Lossy
Compression
18
Geometric
Distortions
20
Signal
Processing
22
D/A-A/D
Conversion
Typical Distortions or Intentional Tampering
26
Transmission
W
24
Corrupted
Watermarked Image or Sound

FIGURE 1

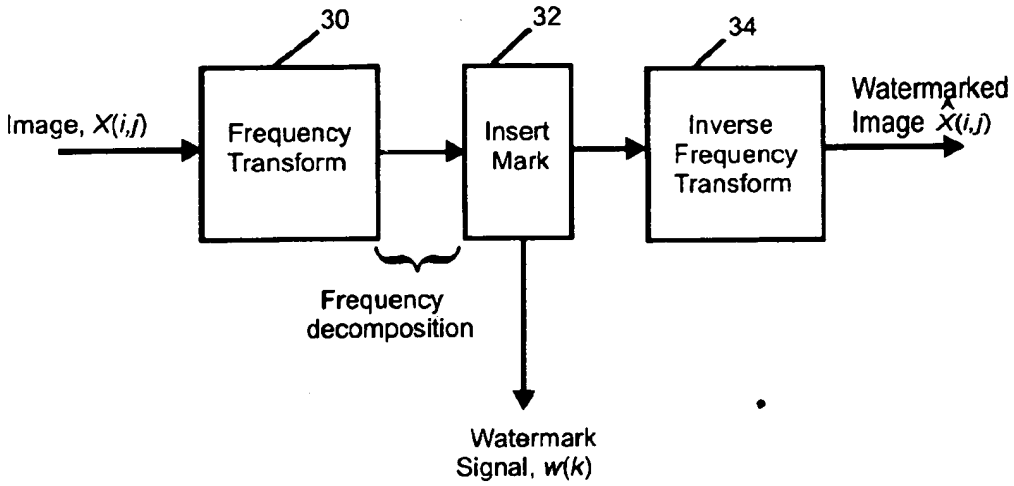


FIGURE 2

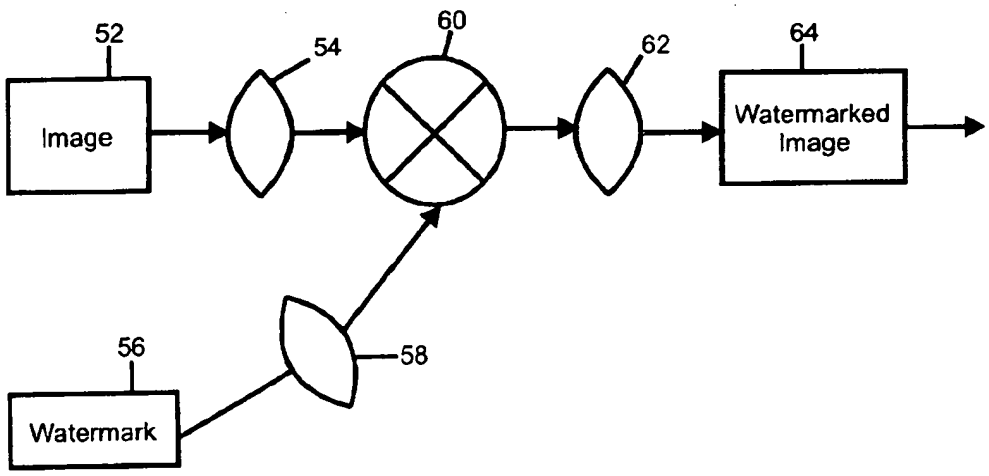


FIGURE 7

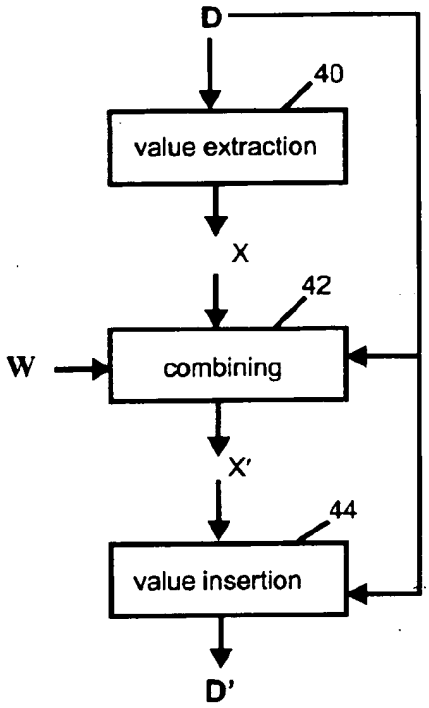


FIGURE 3a

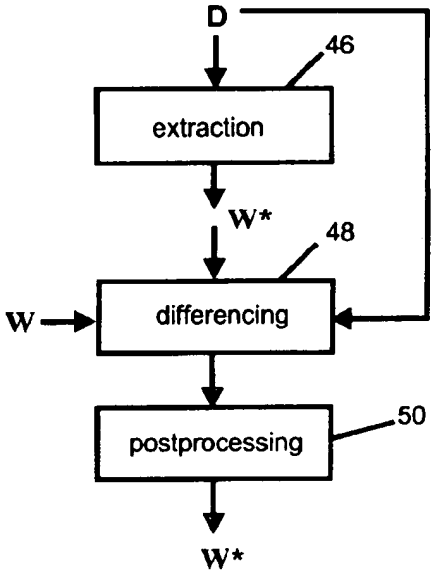


FIGURE 3b

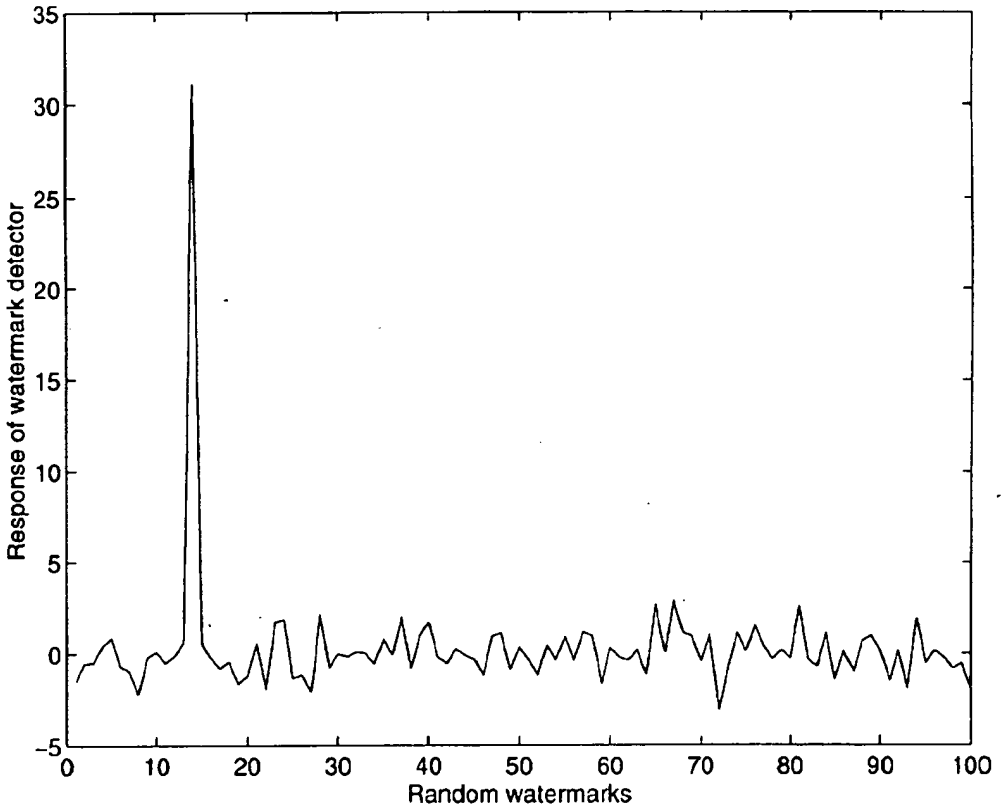

Response of watermark detector
Random watermarks
35
30
25
20
15
10
5
0
-5
0
10
20
30
40
50
60
70
80
90
100

FIGURE 4

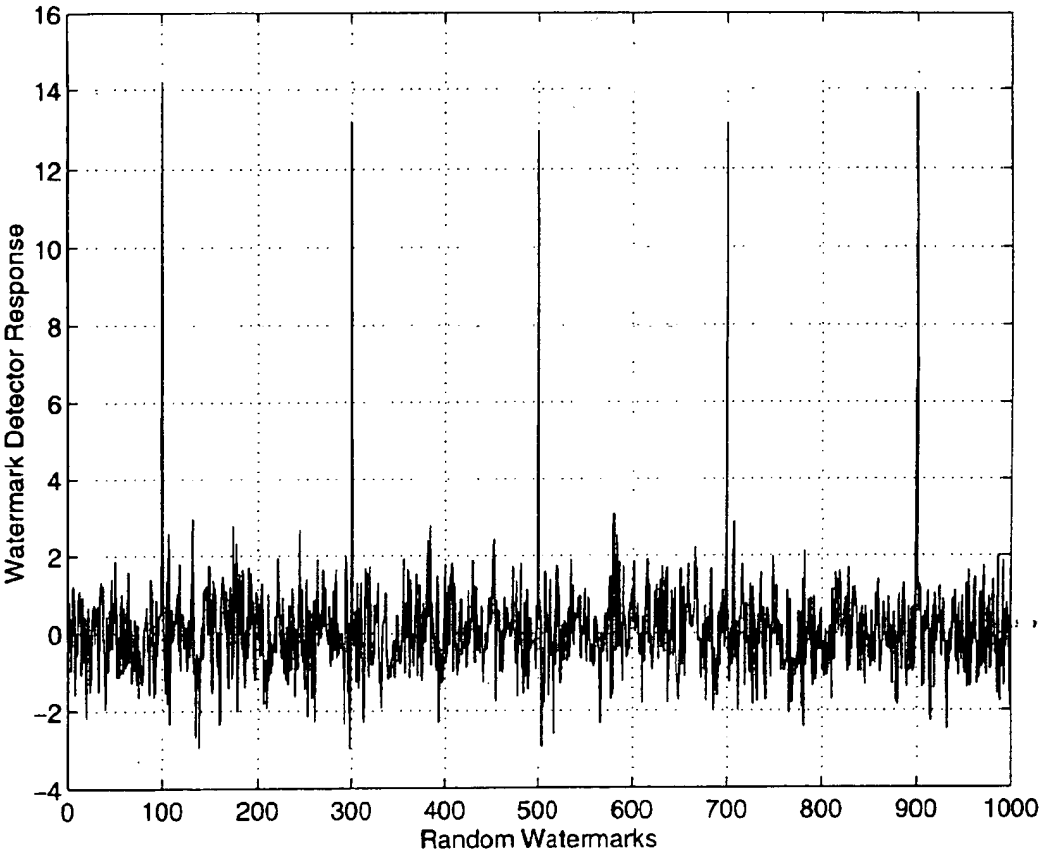

Watermark Detector Response
16
14
12
10
8
6
4
2
0
-2
-4
0
100
200
300
400
500
600
700
800
900
1000
Random Watermarks

FIGURE 5

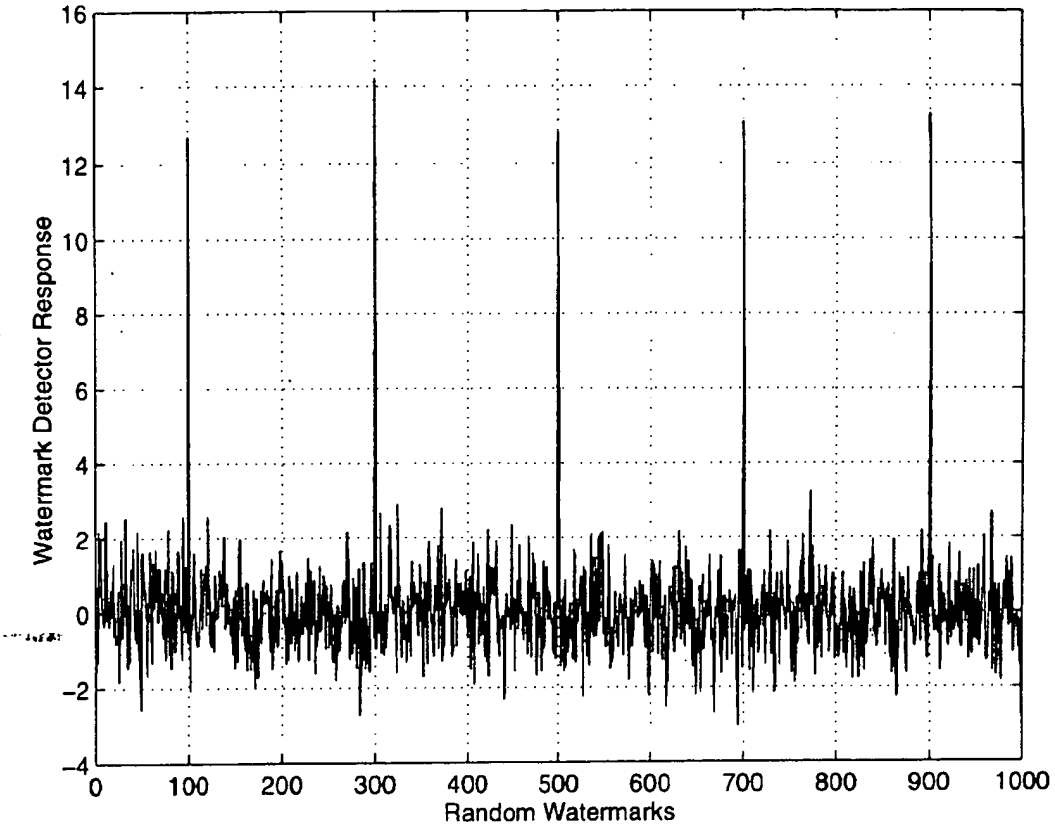

Watermark Detector Response
16
14
12
10
8
6
4
2
0
−2
−4
0
100
200
300
400
500
600
700
800
900
1000
Random Watermarks

FIGURE 6

SECURE SPREAD SPECTRUM WATERMARKING FOR MULTIMEDIA DATA

This application is a continuation of application Ser. No. 08/534,894, filed Sep. 28, 1995, now abandoned.

FIELD OF THE INVENTION

The present invention concerns a method of digital watermarking for use in audio, image, video and multimedia data for the purpose of authenticating copyright ownership, identifying copyright infringers or transmitting a hidden message. Specifically, a watermark is inserted into the perceptually most significant components of a decomposition of the data in a manner so as to be virtually imperceptible. More specifically, a narrow band signal representing the watermark is placed in a wideband channel that is the data.

BACKGROUND OF THE INVENTION

The proliferation of digitized media such as audio, image and video is creating a need for a security system which facilitates the identification of the source of the material. The need manifests itself in terms of copyright enforcement and identification of the source of the material.

Using conventional cryptographic systems permits only valid keyholder access to encrypted data, but once the data is encrypted, it is not possible to maintain records of its subsequent representation or transmission. Conventional cryptography therefore provides minimal protection against data piracy of the type a publisher or owner of data or material is confronted with by unauthorized reproduction or distribution of such data or material.

A digital watermark is intended to complement cryptographic processes. The watermark is a visible or preferably an invisible identification code that is permanently embedded in the data. That is, the watermark remains with the data after any decryption process. As used herein the terms data and material will be understood to refer to audio (speech and music), images (photographs and graphics), video (movies or sequences of images) and multimedia data (combinations of the above categories of materials) or processed or compressed versions thereof. These terms are not intended to refer to ASCII representations of text, but do refer to text represented as an image. A simple example of a watermark is a visible "seal" placed over an image to identify the copyright owner. However, the watermark might also contain additional information, including the identity of the purchaser of the particular copy of the image. An effective watermark should possess the following properties:

1. The watermark should be perceptually invisible or its presence should not interfere with the material being protected.

2. The watermark must be difficult (preferably virtually impossible) to remove from the material without rendering the material useless for its intended purpose. However, if only partial knowledge is known, e.g. the exact location of the watermark within an image is unknown, then attempts to remove or destroy the watermark, for instance by adding noise, should result in severe degradation in data fidelity, rendering the data useless, before the watermark is removed or lost.

3. The watermark should be robust against collusion by multiple individuals who each possess a watermarked copy of the data. That is, the watermark should be robust to the combining of copies of the same data set to destroy the watermarks. Also, it must not be possible for colluders to combine each of their images to generate a different valid watermark.

4. The watermark should still be retrievable if common signal processing operations are applied to the data. These operations include, but are not limited to digital-to-analog and analog-to-digital conversion, resampling, requantization (including dithering and recompression) and common signal enhancements to image contrast and color, or audio bass and treble for example. The watermarks in image and video data should be immune from geometric image operations such as rotation, translation, cropping and scaling.

5. The same digital watermark method or algorithm should be applicable to each of the different media under consideration. This is particularly useful in watermarking of multimedia material. Moreover, this feature is conducive to the implementation of video and image/video watermarking using common hardware.

6. Retrieval of the watermark should unambiguously identify the owner. Moreover, the accuracy of the owner identification should degrade gracefully during attack. Several previous digital watermarking methods have been proposed. L. F. Turner in patent number W089/08915 entitled "Digital Data Security System" proposed a method for inserting an identification string into a digital audio signal by substituting the "insignificant" bits of randomly selected audio samples with the bits of an identification code. Bits are deemed "insignificant" if their alteration is inaudible. Such a system is also appropriate for two dimensional data such as images, as discussed in an article by R. G. Van Schyndel et al entitled "A digital watermark" in Intl. Conf. on Image Processing, vol 2, Pages 86–90, 1994. The Turner method may easily be circumvented. For example, if it is known that the algorithm only affects the least significant two bits of a word, then it is possible to randomly flip all such bits, thereby destroying any existing identification code.

An article entitled "Assuring Ownership Rights for Digital Images" by G. Caronni, in Proc. Reliable IT Systems, VIS '95, 1995 suggests adding tags—small geometric patterns-to-digitized images at brightness levels that are imperceptible. While the idea of hiding a spatial watermark in an image is fundamentally sound, this scheme is susceptible to attack by filtering and redigitization. The fainter such watermarks are, the more susceptible they are to such attacks and geometric shapes provide only a limited alphabet with which to encode information. Moreover, the scheme is not applicable to audio data and may not be robust to common geometric distortions, especially cropping. J. Brassil et al in an article entitled "Electronic Marking and Identification Techniques to Discourage Document Copying" in Proc. of Infocom 94, pp 1278–1287, 1994 propose three methods appropriate for document images in which text is common. Digital watermarks are coded by: (1)vertically shifting text lines, (2) horizontally shifting words, or (3) altering text features such as the vertical endlines of individual characters. Unfortunately, all three proposals are easily defeated, as discussed by the authors. Moreover, these techniques are restricted exclusively to images containing text.

An article by K. Tanaka et al entitled "Embedding Secret Information into a Dithered Multi-level Image" in IEEE Military Comm. Conf., pp216–220, 1990 and K. Mitsui et al in an article entitled "Video-Steganography" in IMA Intellectual Property Proc., vl, pp187–206, 1994, describe several watermarking schemes that rely on embedding watermarks that resemble quantization noise. Their ideas hinge on the notion that quantization noise is typically imperceptible to viewers. Their first scheme injects a watermark into an image by using a predetermined data stream to guide level selection in a predictive quantizer. The data stream is chosen so that the resulting watermark looks like quantization noise.

A variation of this scheme is also presented, where a watermark in the form of a dithering matrix is used to dither an image in a certain way. There are several drawbacks to these schemes. The most important is that they are susceptible to signal processing, especially requantization, and geometric attacks such as cropping. Furthermore, they degrade an image in the same way that predictive coding and dithering can.

In Tanaka et al, the authors also propose a scheme for watermarking facsimile data. This scheme shortens or lengthens certain runs of data in the run length code used to generate the coded fax image. This proposal is susceptible to digital-to-analog and analog-to digital conversions. In particular, randomizing the least significant bit (LSB) of each pixel's intensity will completely alter the resulting run length encoding. Tanaka et al also propose a watermarking method for "color-scaled picture and video sequences". This method applies the same signal transform as JPEG (DCT of 8×8 sub-blocks of an image) and embeds a watermark in the coefficient quantization module. While being compatible with existing transform coders, this scheme is quite susceptible to requantization and filtering and is equivalent to coding the watermark in the least significant bits of the transform coefficients.

In a recent paper, by Macq and Quisquater entitled "Cryptology for Digital TV Broadcasting" in Proc. of the IEEE, 83(6), pp944–957, 1995 there is briefly discussed the issue of watermarking digital images as part of a general survey on cryptography and digital television. The authors provide a description of a procedure to insert a watermark into the least significant bits of pixels located in the vicinity of image contours. Since it relies on modifications of the least significant bits, the watermark is easily destroyed. Further, the method is only applicable to images in that it seeks to insert the watermark into image regions that lie on the edge of contours.

W. Bender et al in article entitled "Techniques for Data Hiding" in Proc. of SPIE, v2420, page 40, July 1995, describe two watermarking schemes. The first is a statistical method called "Patchwork". Patchwork randomly chooses n pairs of image points (a_i, b_i) and increases the brightness at a_i by one unit while correspondingly decreasing the brightness of b_i. The expected value of the sum of the differences of the n pairs of points is claimed to be 2n, provided certain statistical properties of the image are true. In particular, it is assumed that all brightness levels are equally likely, that is, intensities are uniformly distributed. However, in practice, this is very uncommon. Moreover, the scheme may not be robust to randomly jittering the intensity levels by a single unit, and be extremely sensitive to geometric affine transformations.

The second method is called "texture block coding", where a region of random texture pattern found in the image is copied to an area of the image with similar texture. Autocorrelation is then used to recover each texture region. The most significant problem with this technique is that it is only appropriate for images that possess large areas of random texture. The technique could not be used on images of text, for example. Nor is there a direct analog for audio.

In addition to direct work on watermarking images, there are several works of interest in related areas. E. H. Adelson in U.S. Pat. No. 4,939,515 entitled "Digital Signal Encoding and Decoding Apparatus" describes a technique for embedding digital information in an analog signal for the purpose of inserting digital data into an analog TV signal. The analog signal is quantized into one of two disjoint ranges ({0,2,4 . . . }, {1,3,5}, for example) which are selected based on the binary digit to be transmitted. Thus Adelson's method is equivalent to watermark schemes that encode information into the least significant bits of the data or its transform coefficients. Adelson recognizes that the method is susceptible to noise and therefore proposes an alternative scheme wherein a 2×1 Hadamard transform of the digitized analog signal is taken. The differential coefficient of the Hadamard transform is offset by 0 or 1 unit prior to computing the inverse transform. This corresponds to encoding the watermark into the least significant bit of the differential coefficient of the Hadamard transform. It is not clear that this approach would demonstrate enhanced resilience to noise. Furthermore, like all such least significant bit schemes, an attacker can eliminate the watermark by randomization.

U.S. Pat. No. 5,010,405 describes a method of interleaving a standard NTSC signal within an enhanced definition television (EDTV) signal. This is accomplished by analyzing the frequency spectrum of the EDTV signal (larger than that of the NTSC signal) and decomposing it into three sub-bands (L,M,H for low, medium and high frequency respectively). In contrast, the NTSC signal is decomposed into two subbands, L and M. The coefficients, M_k, within the M band are quantized into M levels and the high frequency coefficients, H_k, of the EDTV signal are scaled such that the addition of the H_k signal plus any noise present in the system is less than the minimum separation between quantization levels. Once more, the method relies on modifying least significant bits. Presumably, the mid-range rather than low frequencies were chosen because they are less perceptually significant. In contrast, the method proposed in the present invention modifies the most perceptually significant components of the signal.

Finally, it should be noted that many, if not all, of the prior art protocols are not collusion resistant.

Recently, Digimarc Corporation of Portland, Oreg., has described work referred to as signature technology for use in identifying digital intellectual property. Their method adds or subtracts small random quantities from each pixels. Addition or subtraction is based on comparing a binary mask of N bits with the least significant bit (LSB) of each pixel. If the LSB is equal to the corresponding mask bit, then the random quantity is added, otherwise it is subtracted. The watermark is extracted by first computing the difference between the original and watermarked images and then by examining the sign of the difference, pixel by pixel, to determine if it corresponds to the original sequence of additions/subtractions. The Digimarc technique is not based on direct modifications of the image spectrum and does not make use of perceptual relevance. While the technique appears to be robust, it may be susceptible to constant brightness offsets and to attacks based on exploiting the high degree of local correlation present in an image. For example, randomly switching the position of similar pixels within a local neighborhood may significantly degrade the watermark without damaging the image.

In a paper by Koch, Rindfrey and Zhao entitled "Copyright Protection for Multimedia Data", two general methods for watermarking images are described. The first method partitions an image into 8×8 blocks of pixels and computes the Discrete Cosine Transform (DCT) of each of these blocks. A pseudorandom subset of the blocks is chosen and in each such block a triple of frequencies selected from one of 18 predetermined triples is modified so that their relative strengths encode a 1 or 0 value. The 18 possible triples are composed by selection of three out of eight predetermined frequencies within the 8×8 DCT block. The choice of the

eight frequencies to be altered within the DCT block appears to be based on the belief that middle frequencies have a moderate variance level, i.e., they have similar magnitude. This property is needed in order to allow the relative strength of the frequency triples to be altered without requiring a modification that would be perceptually noticeable. Unlike in the present invention, the set of frequencies is not chosen based on any perceptual significance or relative energy considerations. In addition, because the variance between the eight frequency coefficients is small, one would expect that the technique may be sensitive to noise or distortions. This is supported by the experimental results reported in the Koch et al paper, supra, where it is reported that the "embedded labels are robust against JPEG compression for a quality factor as low as about 50%". In contrast, the method described in accordance with the teachings of the present invention has been demonstrated with compression quality factors as low as 5 percent.

An earlier proposal by Koch and Zhao in a paper entitled "Toward Robust and Hidden Image Copyright Labeling" proposed not triples of frequencies but pairs of frequencies and was again designed specifically for robustness to JPEG compression. Nevertheless, the report states that "a lower quality factor will increase the likelihood that the changes necessary to superimpose the embedded code on the signal will be noticeably visible".

In a second method, proposed by Koch and Zhao, designed for black and white images, no frequency transform is employed. Instead, the selected blocks are modified so that the relative frequency of white and black pixels encodes the final value. Both watermarking procedures are particularly vulnerable to multiple document attacks. To protect against this, Zhao and Koch proposed a distributed 8×8 block of pixels created by randomly sampling 64 pixels from the image. However, the resulting DCT has no relationship to that of the true image. Consequently, one would expect such distributed blocks to be both sensitive to noise and likely to cause noticeable artifacts in the image.

In summary, prior art digital watermarking techniques are not robust and the watermark is easy to remove. In addition, many prior techniques would not survive common signal and geometric distortions

SUMMARY OF THE INVENTION

The present invention overcomes the limitations of the prior art methods by providing a watermarking system that embeds an unique identifier into the perceptually significant components of a decomposition of an image, an audio signal or a video sequence.

Preferably, the decomposition is a spectral frequency decomposition. The watermark is embedded in the data's perceptually significant frequency components. This is because an effective watermark cannot be located in perceptually insignificant regions of image data or in its frequency spectrum, since many common signal or geometric processes affect these components. For example, a watermark located in the high frequency spectral components of an image is easily removed, with minor degradation to the image, by a process that performs low pass filtering. The issue then becomes one of how to insert the watermark into the most significant regions of the data frequency spectrum without the alteration being noticeable to an observer, i.e., a human or a machine feature recognition system. Any spectral component may be altered, provided the alteration is small. However, very small alterations are susceptible to any noise present or intentional distortion.

In order to overcome this problem, the frequency domain of the image data or sound data may be considered as a communication channel, and correspondingly the watermark may be considered as a signal transmitted through the channel. Attacks and intentional signal distortions are thus treated as noise from which the transmitted signal must be immune. Attacks are intentional efforts to remove, delete or otherwise overcome the beneficial aspects of the data watermarking. While the present invention is intended to embed watermarks in data, the same methodology can be applied to sending any type of message through media data.

Instead of encoding the watermark into the least significant components of the data, the present invention considers applying concepts of spread spectrum communication. In spread spectrum communications, a narrowband signal is transmitted over a much larger bandwidth such that the signal energy present in any single frequency is imperceptible. In a similar manner, the watermark is spread over many frequency bins so that the energy in any single bin is small and imperceptible. Since the watermark verification process includes a priori knowledge of the locations and content of the watermarks, it is possible to concentrate these many weak signals into a single signal with a high signal to-noise ratio. Destruction of such a watermark would require noise of high amplitude to be added to every frequency bin.

In accordance with the teachings of the present invention, a watermark is inserted into the perceptually most significant regions of the data decomposition. The watermark itself is designed to appear to be additive random noise and is spread throughout the image. By placing the watermark into the perceptually significant components, it is much more difficult for an attacker to add more noise to the components without adversely affecting the image or other data. It is the fact that the watermark looks like noise and is spread throughout the image or data which makes the present scheme appear to be similar to spread spectrum methods used in communications system.

Spreading the watermark throughout the spectrum of an image ensures a large measure of security against unintentional or intentional attack. First, the location of the watermark is not obvious. Second, frequency regions are selected in a fashion that ensures severe degradation of the original data following any attack on the watermark.

A watermark that is well placed in the frequency domain of an image or a sound track will be practically impossible to see or hear. This will always be the case if the energy in the watermark is sufficiently small in any single frequency coefficient. Moreover, it is possible to increase the energy present in particular frequencies by exploiting knowledge of masking phenomena in the human auditory and visual systems. Perceptual masking refers to any situation where information in certain regions of an image or a sound is occluded by perceptually more prominent information in another part of the image or sound. In digital waveform coding, this frequency domain (and in some cases, time/pixel domain) masking is exploited extensively to achieve low bit rate encoding of data. It is clear that both auditory and visual systems attach more resolution to the high energy, low frequency, spectral regions of an auditory or visual scene. Further, spectrum analysis of images and sounds reveals that most of the information in such data is often located in the low frequency regions.

In addition, particularly for processed or compressed data, perceptually significant need not refer to human perceptual significance, but may refer instead to machine perceptual significance, for instance, machine feature recognition.

To meet these requirements, a watermark is proposed whose structure comprises a large quantity, for instance 1000, of randomly generated numbers with a normal distribution having zero mean and unity variance. A binary watermark is not chosen because it is much less robust to attacks based on collusion of several independently watermarked copies of an image. However, generally, the watermark might have arbitrary structure, both deterministic and/or random, and including uniform distributions. The length of the proposed watermark is variable and can be adjusted to suit the characteristics of the data. For example, longer watermarks might be used for images that are especially sensitive to large modifications of its spectral coefficients, thus requiring weaker scaling factors for individual components.

The watermark is then placed in components of the image spectrum. These components may be chosen based on an analysis of those components which are most vulnerable to attack and/or which are most perceptually significant. This ensures that the watermark remains with the image even after common signal and geometric distortions. Modification of these spectral components results in severe image degradation long before the watermark itself is destroyed. Of course, to insert the watermark, it is necessary to alter these very same coefficients. However, each modification can be extremely small and, in a manner similar to spread spectrum communication, a strong narrowband watermark may be distributed over a much broader image (channel) spectrum. Conceptually, detection of the watermark then proceeds by adding all of these very small signals, whose locations are only known to the copyright owner, and concentrating the watermark into a signal with high signal-to-noise ratio. Because the location of the watermark is only known to the copyright holder, an attacker would have to add very much more noise energy to each spectral coefficient in order to be confident of removing the watermark. However, this process would destroy the image.

Preferably, a predetermined number of the largest coefficients of the DCT (discrete cosine transform) (excluding the DC term) are used. However, the choice of the DCT is not critical to the algorithm and other spectral transforms, including wavelet type decompositions are also possible. In fact, use of the FFT rather than DCT is preferable from a computational perspective.

The invention will be more clearly understood when the following description is read in conjunction with the accompanying drawing.

BRIEF DESCRIPTION OF THE DRAWING

FIG. **1** is a schematic representation of typical common processing operations to which data could be subjected;

FIG. **2** is a schematic representation of a preferred system for immersing a watermark into an image;

FIGS. **3***a* and **3***b* are flow charts of the encoding and decoding of watermarks;

FIG. **4** is a graph of the responses of the watermark detector to random watermarks;

FIG. **5** is a graph of the response of the watermark detector to random watermarks for an image which is successively watermarked five times;

FIG. **6** is a graph of the response of the watermark detector to random watermarks where five images, each having a different watermark, and averaged together; and

FIG. **7** is a schematic diagram of an optical embodiment of the present invention

DETAILED DESCRIPTION

In order to better understand the advantages of the invention, the preferred embodiment of a frequency spectrum based watermarking system will be described. It is instructive to examine the processing stages that image (or sound) data may undergo in the copying process and to consider the effect that such processing stages can have on the data. Referring to FIG. **1**, a watermarked image or sound data **10** is transmitted **12** to undergo typical distortion or intentional tampering **14**. Such distortions or tampering includes lossy compression **16**, geometric distortion **18**, signal processing **20** and D/A and A/D conversion **22**. After undergoing distortion or tampering, corrupted watermarked image or sound data **24** is transmitted **26**. The process of "transmission" refers to the application of any source or channel code and/or of encryption techniques to the data. While most transmission steps are information lossless, many compression schemes (e.g., JPEG, MPEG, etc.) may potentially degrade the quality of the data through irretrievable loss of data. In general, a watermarking method should be resilient to any distortions introduced by transmission or compression algorithms.

Lossy compression **16** is an operation that usually eliminates perceptually irrelevant components of image or sound data. In order to preserve a watermark when undergoing lossy compression, the watermark is located in a perceptually significant region of the data. Most processing of this type occurs in the frequency domain. Data loss usually occurs in the high frequency components. Thus, the watermark must be placed in the significant frequency component of the image (or sound) data spectrum to minimize the adverse affects of lossy compression.

After receipt, an image may encounter many common transformations that are broadly categorized as geometric distortions or signal distortions. Geometric distortions 18 are specific to image and video data, and include such operations as rotation, translation, scaling and cropping. By manually determining a minimum of four or nine corresponding points between the original and the distorted watermark, it is possible to remove any two or three dimensional affine transformation. However, an affine scaling (shrinking) of the image results in a loss of data in the high frequency spectral regions of the image. Cropping, or the cutting out and removal of portions of an image, also results in irretrievable loss of data. Cropping may be a serious threat to any spatially based watermark but is less likely to affect a frequency-based scheme.

Common signal distortions include digital-to-analog and analog-to-digital conversion 22, resampling, requantization, including dithering and recompression, and common signal enhancements to image contrast and/or color, and audio frequency equalization. Many of these distortions are non-linear, and it is difficult to analyze their effect in either a spatial or frequency based method..However, the fact that the original image is known allows many signal transformations to be undone, at least approximately. For example, histogram equalization, a common non-linear contrast enhancement method, may be substantially removed by histogram specification or dynamic histogram warping techniques.

Finally, the copied image may not remain in digital form. Instead, it is likely to be printed or an analog recording made (analog audio or video tape). These reproductions introduce additional degradation into the image data that a watermarking scheme must be robust to.

Tampering (or attack) refers to any intentional attempt to remove the watermark, or corrupt it beyond recognition. The

watermark must not only be resistant to the inadvertent application of distortions. It must also be immune to intentional manipulation by malicious parties. These manipulations can include combinations of distortions, and can also include collusion and forgery attacks.

FIG. **2** shows a preferred system for inserting a watermark into an image in the frequency domain. Image data X(i,j) assumed to be in digital form, or alternatively data in other formats such as photographs, paintings or the like, that have been previously digitized by well-known methods, is subject to a frequency transformation **30**, such as the Fourier transform. A watermark signal W (k) is inserted into the frequency spectrum components of the transformed image data **32** applying the techniques described below. The frequency spectrum image data including the watermark signal is subjected to an inverse frequency transform **34**, resulting in watermarked image data $\hat{X}(i,j)$, which may remain in digital form or be printed as an analog representation by well-known methods.

After applying a frequency transformation to the image data **30**, a perceptual mask is computed that highlights prominent regions in the frequency spectrum capable of supporting the watermark without overly affecting perceptual fidelity. This may be performed by using knowledge of the perceptual significance of each frequency in the spectrum, as discussed earlier, or simply by ranking the frequencies based on their energy. The latter method was used in experiments described below.

In general, it is desired to place the watermark in regions of the spectrum that are least affected by common signal distortions and are most significant to image quality as perceived by a viewer, such that significant modification would destroy the image fidelity. In practice, these regions could be experimentally identified by applying common signal distortions to images and examining which frequencies are most affected, and by psychophysical studies to identify how much each component may be modified before significant changes in the image are perceivable.

The watermark signal is then inserted into these prominent regions in a way that makes any tampering create visible (or audible) defects in the data. The requirements of the watermark mentioned above and the distortions common to copying provide constraints on the design of an electronic watermark.

In order to better understand the watermarking method, reference is made to FIGS. **3**(*a*) and **3**(*b*) where from each document D a sequence of values $X=x_1, \ldots, x_n$ is extracted **40** with which a watermark $W=w_1, \ldots, w_n$ is combined **42** to create an adjusted sequence of values $X'=x'_1, \ldots, x'_n$ which is then inserted back **44** into the document in place of values X in order to obtain a watermark document D'. An attack of the document D', or other distortion, will produce a document D*. Having the original document D and the document D*, a possibly corrupted watermark W* is extracted **46** and compared to watermark W **48** for statistical analysis **50**. The values W* are extracted by first extracting a set of values $X^*=x_1^*, \ldots, x_n^*$ from D* (using information about D) and then generating W* from the values X* and the values X.

When combining the values X with the watermark values W in step **42**, scaling parameter α is specified. The scaling parameter α determines the extent to which values W alter values X. Three preferred formulas for computing X' are:

$$x_i'=x_i+\alpha w_i \quad (1)$$

$$x_i'=x_i(1+\alpha w_i) \quad (2)$$

$$x_i'=x_i(e^{\alpha w_i}) \quad (3)$$

Equation 1 is invertible. Equations 2 and 3 are invertible when $x_i \neq 0$. Therefore, given X* it is possible to compute the inverse function necessary to derive W* from X and X*.

Equation 1 is not the preferred formula when the values x_i vary over a wide range. For example, if $x_i=10^6$ then adding 100 may be insufficient to establish a watermark, but if $x_i=10$, then adding 100 will unacceptably distort the value. Insertion methods using equations 2 and 3 are more robust when encountering such a wide range of values x_i. It will also be observed that equation 2 and 3 yield similar results when αw_i is small. Moreover, when x_i is positive, equation 3 is equivalent to $\ln(x_i')=\ln(x_i)+\alpha x_i$ and may be considered as an application of equation 1 when natural logarithms of the original values are used. For example, if $|w_i| \leq 1$ and $\alpha=0.01$, then using Equation (2) guarantees that the spectral coefficient will change by no more than 1%.

For certain applications, a single scaling parameter α may not be best for combining all values of x_i. Therefore, multiple scaling parameters $\alpha_i, \ldots, \alpha_n$ can be used with revised equations 1 to 3 such as $x_i'=x_i(1+\alpha_i w_i)$. The values of α_i serve as a relative measure of how much x_i must be altered to change the perceptual quality of the document. A large value for α_i means that it is possible to alter x_i by a large amount without perceptually degrading the document.

A method for selecting the multiple scaling values is based upon certain general assumptions. For example, equation 2 is a special case of the generalized equation 1, $(x_i'=x_i+\alpha_i x_i)$, for $\alpha_i=\alpha x_i$. That is, equation 2 makes the reasonable assumption that a large value of x_i is less sensitive to additive alteration that a small value of x_i.

Generally, the sensitivity of the image to different values of α_i is unknown. A method of empirically estimating the sensitivities is to determine the distortion caused by a number of attacks on the original image. For example, it is possible to compute a degraded image D* from D, extract the corresponding values $x_1^*, \ldots, x_n^*$ and select α_i to be proportional to the deviation $|x_i^*-x_i|$. For greater robustness, it is possible to try other forms of distortion and make α_i proportional to the average value of $|x_i^*-x_i|$. Instead of using the average distortion, it is possible to use the median or maximum deviation.

Alternatively, it is possible to combine the empirical approach with general global assumptions regarding the sensitivity of the values. For example, it might be required that $\alpha_i \geq \alpha_j$ whenever $x_i \geq x_j$. This can be combined with the empirical approach by setting α_i according to

$$\alpha_i \sim \max_{|j|v_j \leq v_i|} |v_j' - v_j|$$

A more sophisticated approach is to weaken the monotonicity constraint to be robust against occasional outliers.

The length of the watermark, n, determines the degree to which the watermark is spread among the relevant components of the image data. As the size of the watermark increases, so does the number of altered spectral components, and the extent to which each component need be altered decreases for the same resilience to noise. Consider watermarks of the form $x_i'=x_i+\alpha w_i$ and a white noise attack by $x_i''=x_i'+r_i$, where r_i are chosen according to independent normal distributions with standard deviation σ. It is possible to recover the watermark when α is proportional to $\sigma/\sqrt{n}$. That is, quadrupling the number of components can halve the magnitude of the watermark placed into each component. The sum of the squares of the deviations remains essentially unchanged.

In general, a watermark comprises an arbitrary sequence of real numbers $W=w_1, \ldots, w_n$. In practice, each value w_i may be chosen independently from a normal distribution N(0,1), where $N(\mu, \sigma^2)$ with mean μ and variance σ^2 or of a uniform distribution from {1,-1} or {0,1}.

It is highly unlikely that the extracted mark W* will be identical to the original watermark W. Even the act of requantizing the watermarked document for transmission will cause W* to deviate from W. A preferred measure of the similarity of W and W* is

$$sim(W, W^*) = \frac{W^* \cdot W}{\sqrt{W^* \cdot W^*}} \quad (4)$$

Large values of sim (W,W*) are significant in view of the following analysis. Assume that the authors of document D* had no access to W (either through the seller or through a watermarked document). Then for whatever value of W* is obtained, the conditional distribution on w_i will be independently distributed according to N(0,1). In this case,

$$N\left(0, \sum_{i=1}^{n} x_i^{*2}\right) = N(0, W^* \cdot W^*).$$

Thus, sim(W,W*) is distributed according to N(0,1). Then, one may apply the standard significance tests for the normal distribution. For example, if D* is chosen independently from W, then it is very unlikely that sim(W,W*)>5. Note that somewhat higher values of sim (W,W*) may be needed when a large number of watermarks are on file. The above analysis required only the independence of W from W*, and did not rely on any specific properties of W* itself. This fact provides further flexibility when preprocessing W*.

The extracted watermark W* may be extracted in several ways to potentially enhance the ability to extract a watermark. For example, experiments on images encountered instances where the average value of W*, denoted $E_i(W^*)$, differed substantially from 0, due to the effects of a dithering procedure. While this artifact could be easily eliminated as part of the extraction process, it provides a motivation for postprocessing extracted watermarks. As a result, it was discovered that the simple transformation $w_i^* \leftarrow w_i^* - E_i(W^*)$ yielded superior values of sim (W,W*). The improved performance resulted from the decreased value of W*. W*; the value of W*. W was only slightly affected.

In experiments it was frequently observed that w_i^* could be greatly distorted for some values of i. One postprocessing option is to simply ignore such values, setting them to 0. That is,

$$w_i^* \leftarrow \begin{cases} w_i^* & \text{if } |w_i^*| > \text{tolerance} \\ 0 & \text{otherwise} \end{cases}$$

The goal of such a transformation is to lower W*·W*. A less abrupt version of this approach is to normalize the W* values to be either −1,0 or 1, by

$$w_i^* \leftarrow \text{sign}(w_i^* - E_i(W^*)).$$

This transformation can have a dramatic effect on the statistical significance of the result. Other robust statistical techniques could also be used to suppress outlier effects.

In principle, any frequency domain transform can be used. In the scheme described below, a Fourier domain method is used, but the use of wavelet based schemes are also useable as a variation. In terms of selecting frequency regions of the transform, it is possible to use models for the perceptual system under consideration.

Frequency analysis may be performed by a wavelet or sub-band transform where the signal is divided into sub-bands by means of a wavelet or multi-resolution transform. The sub-bands need not be uniformly spaced. Each sub-band may be thought of as representing a frequency region in the domain corresponding to a sub-region of the frequency range of the signal. The watermark is then inserted into the sub-regions.

For audio data, a sliding "window" moves along the signal data and the frequency transform (DCT, FFT, etc.) is taken of the sample in the window. This process enables the capture of meaningful information of a signal that is time varying in nature.

Each coefficient in the frequency domain is assumed to have a perceptual capacity. That is, it can support the insertion of additional information without any (or with minimal) impact to the perceptual fidelity of the data.

In order to place a length L watermark into an N×N image, the N×N FFT (or DCT) of the image is computed and the watermark is placed into the L highest magnitude coefficients of the transform matrix, excluding the DC component. More generally, L randomly chosen coefficients could be chosen from the M, M≧L most perceptually significant coefficients of the transform. For most images, these coefficients will be the ones corresponding to the low frequencies. The purpose of placing the watermark in these locations is because significant tampering with these frequencies will destroy the image fidelity or perceived quality well before the watermark is destroyed.

The FFT provides perceptually similar results to the DCT. This is different than the case of transform coding, where the DCT is preferred to the FFT due to its spectral properties. The DCT tends to have less high frequency information than that the FFT, and places most of the image information in the low frequency regions, making it preferable in situations where data need to be eliminated. In the case of watermarking, image data is preserved, and nothing is eliminated. Thus the FFT is as good as the DCT, and is preferred since it is easier to compute.

In an experiment, a visually imperceptible watermark was intentionally placed in an image. Subsequently, 100 randomly generated watermarks, only one of which corresponded to the correct watermark, were applied to the watermark detector described above. The result, as shown in FIG. 4, was a very strong positive response corresponding to the correct watermark, suggesting that the method results in a very low number of false positive responses and a very low false negative response rate.

In another test, the watermarked image was scaled to half of its original size. In order to recover the watermark, the image was re-scaled to its original size, albeit with loss of detail due to subsampling of the image using low pass spatial filter operations. The response of the watermark detector was well above random chance levels, suggesting that the watermark is robust to geometric distortions. This result was achieved even though 75 percent of the original data was missing from the scaled down image.

In a further experiment, a JPEG encoded version of the image with parameters of 10 percent quality and 0 percent smoothing, resulting in visible distortions, was used. The results of the watermark detector suggest that the method is robust to common encoding distortions. Even using a version of the image with parameters of the 5 percent quality

and 0 percent smoothing, the results were well above that achievable due to random chance.

In experiments using a dithered version of the image, the response of the watermark detector suggested that the method is robust to common encoding distortion. Moreover, more reliable detection is achieved by removing any non-zero mean from the extracted watermark.

In another experiment, the image was clipped, leaving only the central quarter of the image. In order to extract the watermark from the clipped image, the missing portion of the image was replaced with portions from the original unwatermarked image. The watermark detector was able to recover the watermark with a response greater than random. When the non-zero mean was removed, and the elements of the watermark were binarized prior to the comparison with the correct watermark, the detector response was improved. This result is achieved even though 75 percent of the data was removed from the image.

In yet another experiment, the image was printed, photocopied, scanned using a 300 dpi Umax PS-2400x scanner and rescaled to a size of 256x256 pixels. Clearly, the final image suffered from different levels of distortion introduced at each process. High frequency pattern noise was particularly noticeable. When the non-zero mean was removed and only the sign of the elements of the watermark was used, the watermark detector response improved to well above random chance levels.

In still another experiment, the image was subject to five successive watermarking operations. That is, the original image was watermarked, the watermarked image was watermarked, and so forth. The process may be considered another form of attack in which it is clear that significant image degradation occurs if the process is repeated. FIG. **5** shows the response of the watermark detector to 1000 randomly generated watermarks, including the five watermarks present in the image. The five dominant spikes in the graph, indicative of the presence of the five watermarks, show that successive watermarking does not interfere with the process.

The fact that successive watermarking is possible means that the history or pedigree of a document is determinable if successive watermarking is added with each copy.

In a variation of the multiple watermark image, five separately watermarked images were averaged together to simulate simple conclusion attack. FIG. **6** shows the response of the watermark detector to 1000 randomly generated watermarks, including the five watermarks present in the original images. The result is that simple collusion based on averaging is ineffective in defeating the present watermarking system.

The result of the above experiments is that the described system can extract a reliable copy of the watermark from images that have been significantly degraded through several common geometric and signal processing procedures. These procedures include zooming (low pass filtering), cropping, lossy JPEG encoding, dithering, printing, photocopying and subsequent rescanning.

While these experiments were, in fact, conducted using an image, similar results are attainable with text images, audio data and video data, although attention must be paid to the time varying nature of these data.

The above implementation of the watermarking system is an electronic system. Since the basic principle of the invention is the inclusion of a watermark into spectral frequency components of the data, watermarking can be accomplished by other means using, for example, an optical system as shown in FIG. **7**.

In FIG. **7**, data to be watermarked such as an image **52** is passed through a spatial transform lens **54**, such as a Fourier transform lens, the output of which lens is the spatial transform of the image. Concurrently, a watermark image **56** is passed through a second spatial transform lens **58**, the output of which lens is the spatial transfer of the watermark image **56**. The spatial transform from lens **54** and the spatial transform from lens **58** are combined at an optical combiner **60**. The output of the optical combiner **60** is passed through an inverse spatial transform lens **62** from which the watermark image **64** is present. The result is a unique, virtually imperceptible, watermarked image. Similar results are achievable by transmitting video or multimedia signals through the lenses in the manner described above.

While there have been described and illustrated spread spectrum watermarking of data and variations and modifications thereof, it will be apparent to those skilled in the art that further variations and modifications are possible without deviating from the broad principles and spirit of the present invention which shall be limited solely by the scope of the claims appended hereto.

What is claimed is:

1. A method of inserting a watermark into data comprising the steps of:

obtaining a spectral decomposition of data to be watermarked which data is a representation of humanly perceivable material;

inserting a watermark into the perceptually significant components of the decomposition of data: and

applying an inverse transform to the decomposition of data with the watermark for generating watermarked data.

2. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises image data.

3. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises video data.

4. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises audio data.

5. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises multimedia data.

6. A method of inserting a watermark into data as set forth in claim **1**, where said obtaining a spectral decomposition of data is selected from the group consisting of Fourier transformation, discrete cosine transformation, Hadamard transformation, and wavelet, multi-resolution, sub-band method.

7. A method of inserting a watermark into data as set forth in claim **6**, where said inserting a watermark inserts watermark values where addition of additional signal into a perceptually significant component affects the perceived quality of the data.

8. A method of inserting a watermark into data as set forth in claim **1**, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

9. The method of inserting a watermark into data as set forth in claim **8**, wheye watermark values include associated scaling parameters.

10. A method of inserting a watermark into data as set forth in claim **9**, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

11. A method of inserting a watermark into data as set forth in claim 8, where the watermark values are chosen according to a random distribution.

12. A method of inserting a watermark into data comprising the steps of:

extracting values of perceptually significant components of a spectral decomposition of data which data is a representation of human perceivable material;

combining watermark values with the extracted values to create adjusted values; and

inserting the adjusted values into the data in place of the extracted values to produce watermarked data.

13. The method of inserting a watermark into data as set forth in claim 12, where watermark values include associated scaling parameters.

14. A method of inserting a watermark into data as set forth in claim 13, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

15. A method of inserting a watermark into data as set forth in claim 12, where the watermark values are chosen according to a random distribution.

16. A method of inserting a watermark into data as set forth in claim 12, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

17. The method of inserting a watermark into data as set forth in claim 16, where watermark values include associated scaling parameters.

18. A method of inserting a watermark into data as set forth in claim 12, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

19. A method of inserting a watermark into data as set forth in claim 16, where the watermark values are chosen according to a random distribution.

20. A method of inserting a watermark into data as set forth in claim 16, further comprising the step of preprocessing distorted or tampered watermarked data before said comparing data.

21. A method of inserting a watermark into data as set forth in claim 20, where said distorted or tampered watermarked data is clipped data and said preprocessing comprises replacing missing portions of the data with corresponding portions from original unwatermarked data.

22. A method of inserting a watermark into data as set forth in claim 12, where said combining watermark values sequentially combines watermark values for a plurality of watermarks.

23. A system for inserting a watermark into data comprising:

providing image data;

providing watermark data;

first transform lens for transforming image data passing therethrough into transformed image data;

second transform lens for transforming watermark data passing therethrough into transformed watermark data;

optical combiner for combining the transformed image data and the transformed watermark data to form transformed watermarked data; and

inverse transform lens for forming watermarked data by inverse transformation of transformed watermarked data.

24. A system for inserting a watermark into data as set forth in claim 23, where said first transform lens and said second transform lens are Fourier transform lenses and said inverse transform lens is an inverse Fourier transform lens.

25. A method of inserting a watermark into data comprising the steps of:

providing a medium containing data;

obtaining a spectral decomposition of data to be watermarked;

inserting a watermark into the perceptually significant components of the decomposition of data; and

applying an inverse transform to the decomposition of data with the watermark to generate watermarked data.

26. A method of inserting a watermark into data as set forth in claim 25, where said data comprises image data.

27. A method of inserting a watermark into data as set forth in claim 25, where said data comprises video data.

28. A method of inserting a watermark into data as set forth in claim 25, where said data comprises audio data.

29. A method of inserting a watermark into data as set forth in claim 25, where said data comprises multimedia data.

30. A method of inserting a watermark into data as set forth in claim 25, where said obtaining a spectral decomposition of data is selected from the group consisting of Fourier transformation, discrete cosine transformation, Hadamard transformation, and wavelet, multi-resolution, subband method.

31. A method of inserting a watermark into data as set forth in claim 30, where said inserting a watermark inserts watermark values where addition of additional signal into a perceptually significant component affects the perceived quality of the data.

32. A method of inserting a watermark into data as set forth in claim 25, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

33. The method of inserting a watermark into data as set forth in claim 32, where watermark values include associated scaling parameters.

34. A method of inserting a watermark into data as set forth in claim 33, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

35. A method of inserting a watermark into data as set forth in claim 32, where the watermark values are chosen according to a random distribution.

36. A method of inserting a watermark into data comprising the steps of:

providing a medium containing data;

extracting values of perceptually significant components of a spectral decomposition of the data;

combining watermark values with the extracted values to create adjusted values; and

inserting the adjusted values into the data in place of the extracted values to produce watermarked data.

37. The method of inserting a watermark into data as set forth in claim 36, where watermark values include associated scaling parameters.

38. A method of inserting a watermark into data as set forth in claim 37, where scaling parameters are selected such

that adding additional watermark value affects the perceived quality of the data.

39. A method of inserting a watermark into data as set forth in claim 36, where the watermark values are chosen according to a random distribution.

40. A method of inserting a watermark into data as set forth in claim 36, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

41. The method of inserting a watermark into data as set forth in claim 40, where watermark values include associated scaling parameters.

42. A method of inserting a watermark into data as set forth in claim 41, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

43. A method of inserting a watermark into data as set forth in claim 40, where the watermark values are chosen according to a random distribution.

44. A method of inserting a watermark into data as set forth in claim 40, further comprising the step of preprocessing distorted or tampered watermarked data before said comparing data.

45. A method of inserting a watermark into data as set forth in claim 44, where said distorted or tampered watermarked data is clipped data and said preprocessing comprises replacing missing portions of the data with corresponding portions from original unwatermarked data.

46. A method of inserting a watermark into data as set forth in claim 36, where said combining watermark values sequentially combines watermark values for a plurality of watermarks.

* * * * *

US005915027A

United States Patent [19]

Cox et al.

[11] Patent Number: 5,915,027

[45] Date of Patent: Jun. 22, 1999

[54] DIGITAL WATERMARKING

[75] Inventors: Ingemar J. Cox, Lawrenceville, N.J.; Matthew L. Miller, Vilnius, Lithuania; Kazuyoshi Tanaka; Yutaka Wakasu, both of Tokyo, Japan

[73] Assignees: NEC Research Institute, Princeton, N.J.; NEC Corporation, Tokyo, Japan

[21] Appl. No.: 08/746,022

[22] Filed: Nov. 5, 1996

[51] Int. Cl.[6] H04L 9/02

[52] U.S. Cl. 380/54

[58] Field of Search 388/28, 51, 54

[56] References Cited

U.S. PATENT DOCUMENTS

4,939,515	7/1990	Adelson	341/51
5,319,735	6/1994	Preuss et al.	395/2.14
5,530,751	6/1996	Morris	380/4
5,530,759	6/1996	Braudaway et al.	380/54
5,568,570	10/1996	Rabbani	382/238
5,613,004	3/1997	Cooperman et al.	380/28
5,636,292	6/1997	Rhoads	382/232
5,646,997	7/1997	Barton	380/23
5,659,726	8/1997	Sanford, II	395/612
5,687,236	11/1997	Moskowitz et al.	380/28
5,734,752	3/1998	Knox	380/54

FOREIGN PATENT DOCUMENTS

0690595	1/1995	European Pat. Off. .	
2196167	4/1988	United Kingdom .	
8908915	9/1989	WIPO .	
9520291	7/1995	WIPO .	
9621290	7/1996	WIPO	H04H 1/00
9625005	8/1996	WIPO	H04H 7/08
9627259	9/1996	WIPO .	

OTHER PUBLICATIONS

R.G. Van Schyndel et al, "A digital watermark," in Intl. Conf. On Image Processing, vol. 2, pp. 86–90, 1994.

G. Caronni, "Assuring Ownership Rights for Digital Images," in Proc. Reliable IT Systems, VIS '95, 1995.

J. Brassil et al, "Electronic Marking and Identification Techniques to Discourage Document Copying," in Proc. Infocom '94, pp. 1278–1287, 1994.

K. Tanaka et al, "Embedding Secret Information into a Dithered Multi-level Image," in IEEE Military Comm. Conf., pp. 216–220, 1990.

K. Mitsui et al, "Video–Steganography: How to Secretly Embed a Signature in a Picture," in IMA Intellectual Property Project Proc., vol. 1, pp. 187–206, 1994.

Macq and Quisquater, "Cryptology for Digital TV Broadcasting," in Proc. of the IEEE, vol. 83, No. 6, pp. 944–957, 1995.

W. Bender et al, "Techniques for data hiding," in Proc. of SPIE, vol. 2420, No. 40, Jul. 1995.

Koch, Rindfrey and Zhao, "Copyright Protection for Multimedia Data," in Proc. of the Int'l Conf. on Digital Media and Electronic Publishing (Leeds, UK, Dec., 6–8 1994).

Koch and Zhao, "Towards Robust and Hidden Image Copyright Labeling," in Proc. of 1995 IEEE Workshop on Nonlinear Signal and Image Processing (Neos Marmaras, Halkidiki, Greece, Jun. 20–22, 1995.

Zhao and Koch, "Embedding Robust Labels Into Images For Copyright Protection," in Proc. Int. Congr. on IPR for Specialized Information, Knowledge and New Technologies (Vienna, Austria), Aug. 21–25, 1995.

"Digital Copyright: Who Owns What?" NewMedia, Sep. 1995, pp. 38–43.

"Publish and Be Robbed?" New Scientist, Feb. 18, 1995, pp. 32–37.

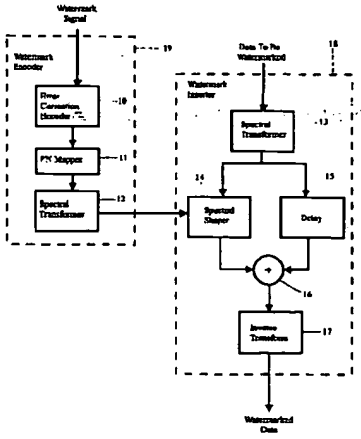
(List continued on next page.)

Primary Examiner—Salvatore Cangialosi
Attorney, Agent, or Firm—Philip J. Feig

[57] ABSTRACT

Digital watermarking of data, including image, video and audio data, is performed by repeatedly inserting the watermark into subregions or subimages of the data. Similarly, the watermark is repeatedly extracted from the subregions of the data.

28 Claims, 8 Drawing Sheets



OTHER PUBLICATIONS

Kohno et al, "Spread Spectrum Access Methods for Wireless Communications," in IEEE Communications Magazine, Jan. 1995, pp. 58–67, 116.

Campana and Quinn, "Spread spectrum communications," in IEEE Potentials, Apr. 1993, pp. 13–16.

Mowbray and Grant, "Wideband coding for uncoordinated multiple access communication," in Electronics & Communication Engineering Journal, Dec. 1992, pp. 351–361.

Digimarc Overview & "Wired" Magazine article (Jul. 1995 issue)—(Jun. 1995).

A.G. Bors et al., "Image Watermarking Using DCT Domain Constraints", Dept. Of Informatics, University of Thessaloniki.

I.J. Cox et al., "Secure Spread Spectrum Watermarking for Multimedia", NEC Research Institute, Technical Report 95–10.

H.S. Stone, "Analysis of Attacks on Image Watermarks with Randomized Coefficients", NEC Research Institute, May 17, 1996.

F.M. Boland et al., "Watermarking Digital Images for Copyright Protection", Image Processing and its Applications, Jul. 4–6, 1995, Conference Publication No. 410, pp. 326–330.

L. Boney et al., "Digital Watermarks for Audio Signals".

Swanson et al., "Transparent Robust Image Watermarking", Proc. IEEE Int. Conf. On Image Proc. 1996.

J.J.K. O Ruanaidh et al., "Phase Watermarking of Digital Images".

I. Pitas, "A Method for Signature Casting on Digital Images".

C.T. Hsu et al., "Hidden Signatures in Images", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

M. Schneider et al., "A Robust Content Based Digital Signature for Image Authentication", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

S. Roche et al., "Multi–Resolution Access Control Algorithm Based on Fractal Coding", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

K. Hirotsugu, "An Image Digital Signature System with ZKIP for the Graph Isomorphism", ICIP 96 Conf. Proc., Sep. 16–19, 1996.

R.B. Wolfgang et al., "A Watermark for Digital Images".

J.J.K. O Ruanaidh et al., "Watermarking Digital Images for Copyright Protection", EVA 96 Florence, pp. 1–7.

T. Aura, "Invisible Communication", Nov. 6, 1995.

D. Kahn, "Information Hiding—An Annotated Bibliography", Macmillan 1967, Library of Congress catalog No. 63–16109.

Craver et al., "Can Invisible Watermarks Resolve Rightful Ownerships?", IBM Research Report.

Podilchuk et al., "Digital Image Watermarking Using Visual Models", Proc. of EI '97, vol. 3016, Feb. 9–14, 1997.

Cox et al., "A review of watermarking and the importance of perceptual modeling", proc. of EI'97, vol. 3016, Feb. 9–14, 1997.

Watson, "DCT quantization matrices visually optimized for individual images", SPIE, vol. 1913, pp. 202–216.

Ahumada, Jr. et al., "Luminance–Model–Based DCT Quantization for Color Image Compression", SPIE, vol. 1666 (1992), pp. 365–374.

Hartung et al., "Digital Watermarking of Raw and Compressed Video", Systems for Video Communication, Oct. 1996, pp. 205–213.

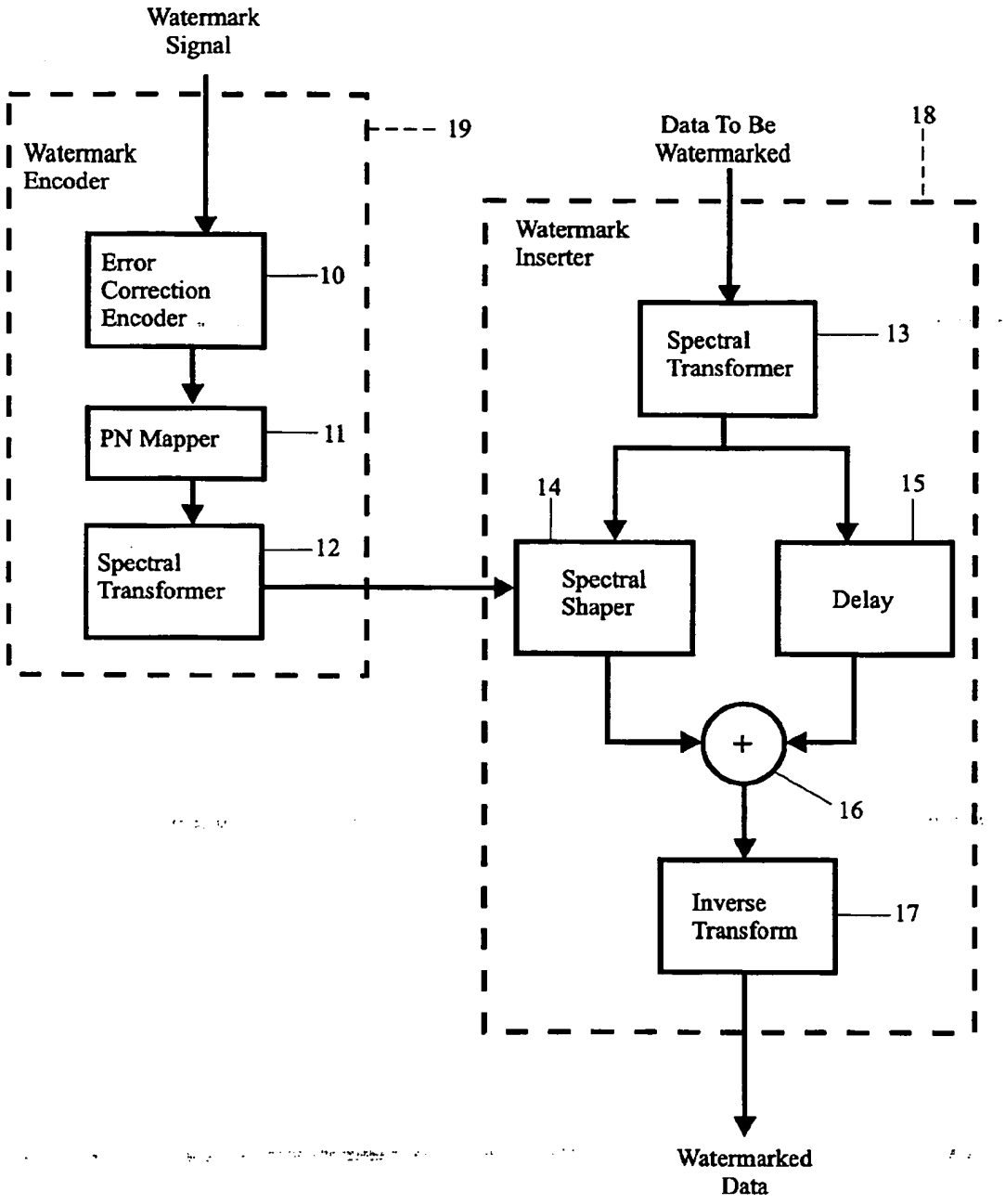

Watermark Signal
Watermark Encoder
19
Error Correction Encoder
10
PN Mapper
11
Spectral Transformer
12
Data To Be Watermarked
18
Watermark Inserter
Spectral Transformer
13
14
15
Spectral Shaper
Delay
+
16
Inverse Transform
17
Watermarked Data

Figure 1

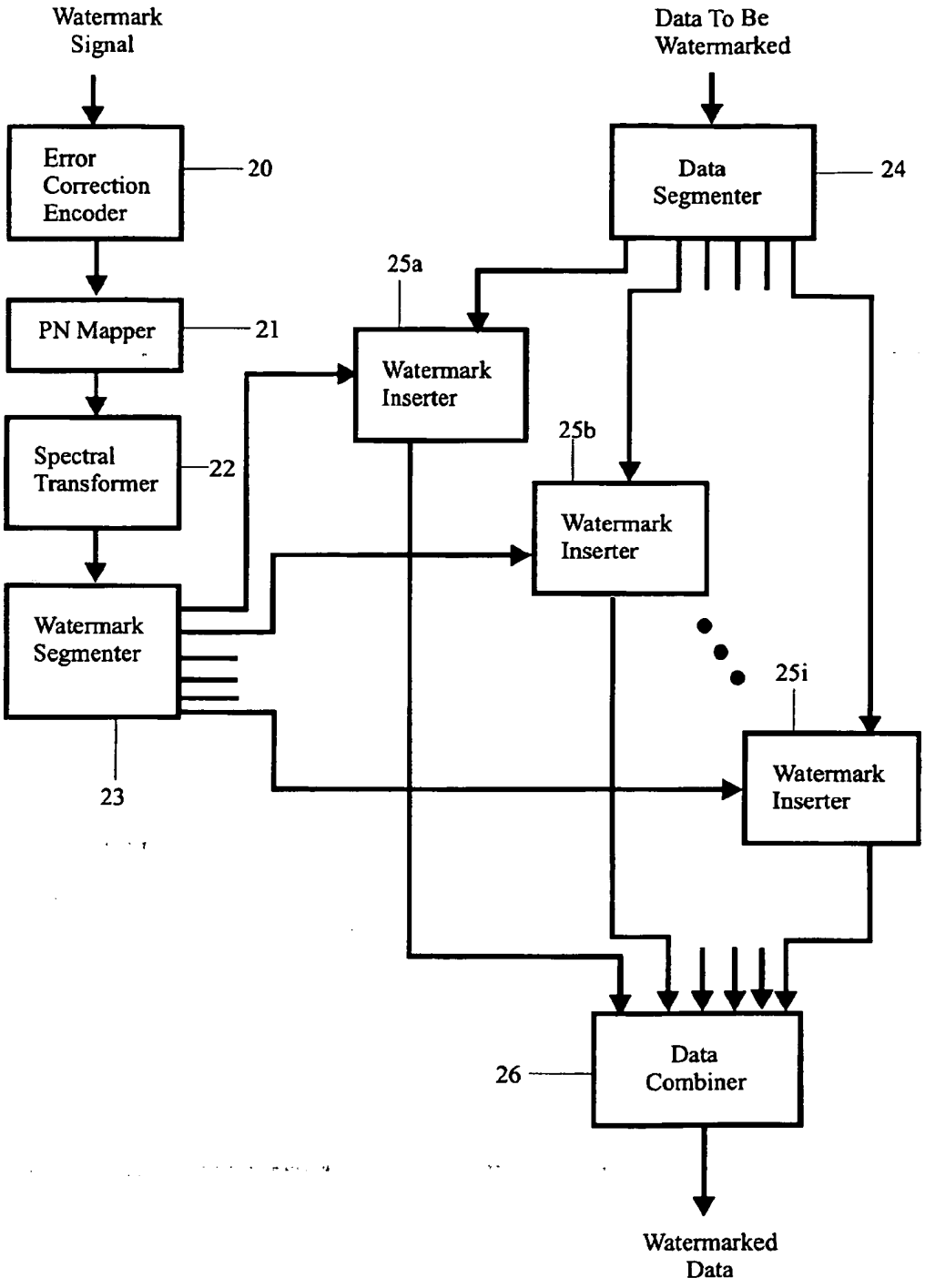

Watermark Signal
Data To Be Watermarked
Error Correction Encoder
20
Data Segmenter
24
25a
PN Mapper
21
Watermark Inserter
25b
Spectral Transformer
22
Watermark Inserter
Watermark Segmenter
25i
Watermark Inserter
23
Data Combiner
26
Watermarked Data

Figure 2

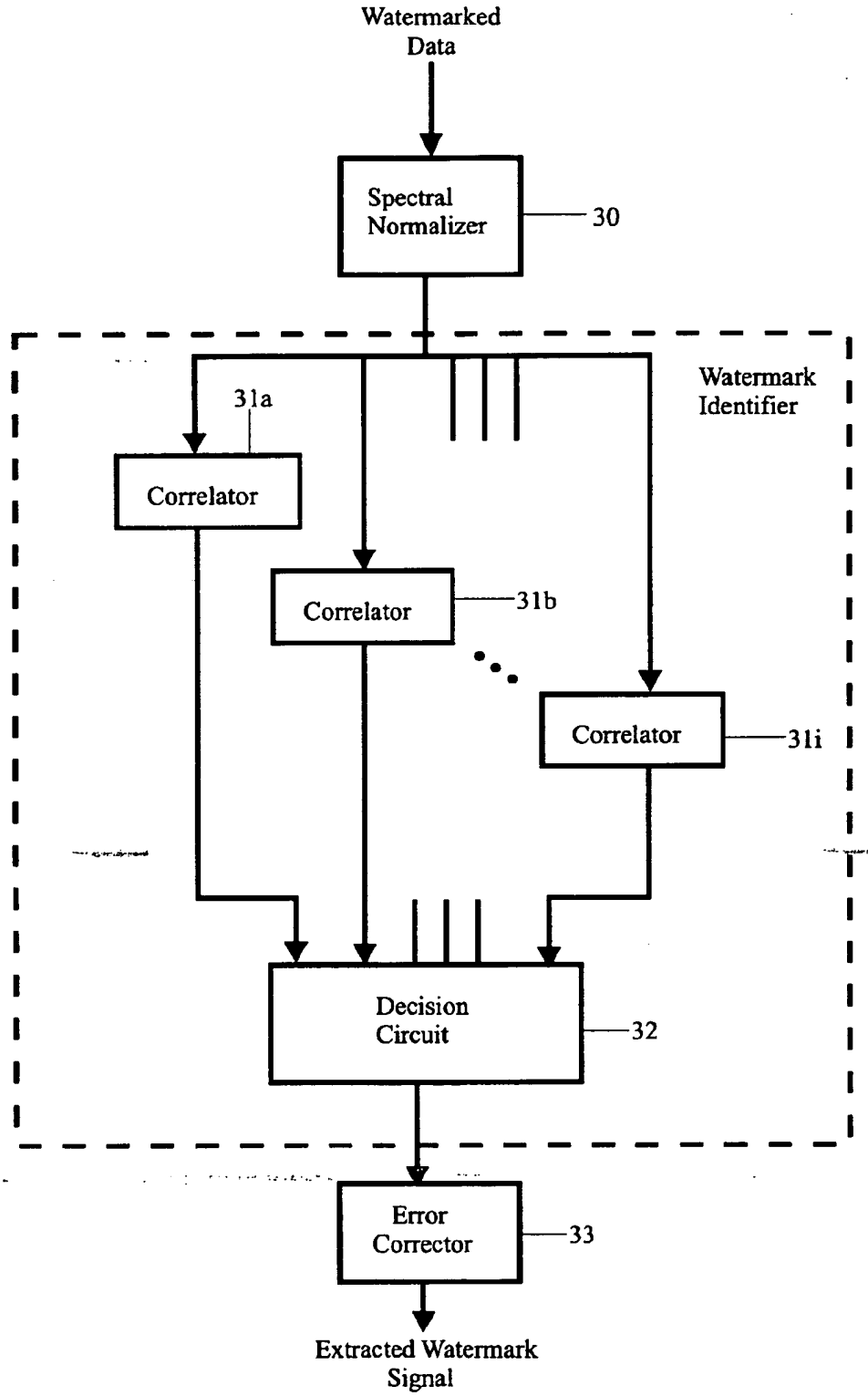

Watermarked Data
Spectral Normalizer
30
Watermark Identifier
31a
Correlator
Correlator
31b
Correlator
31i
Decision Circuit
32
Error Corrector
33
Extracted Watermark Signal

Figure 3

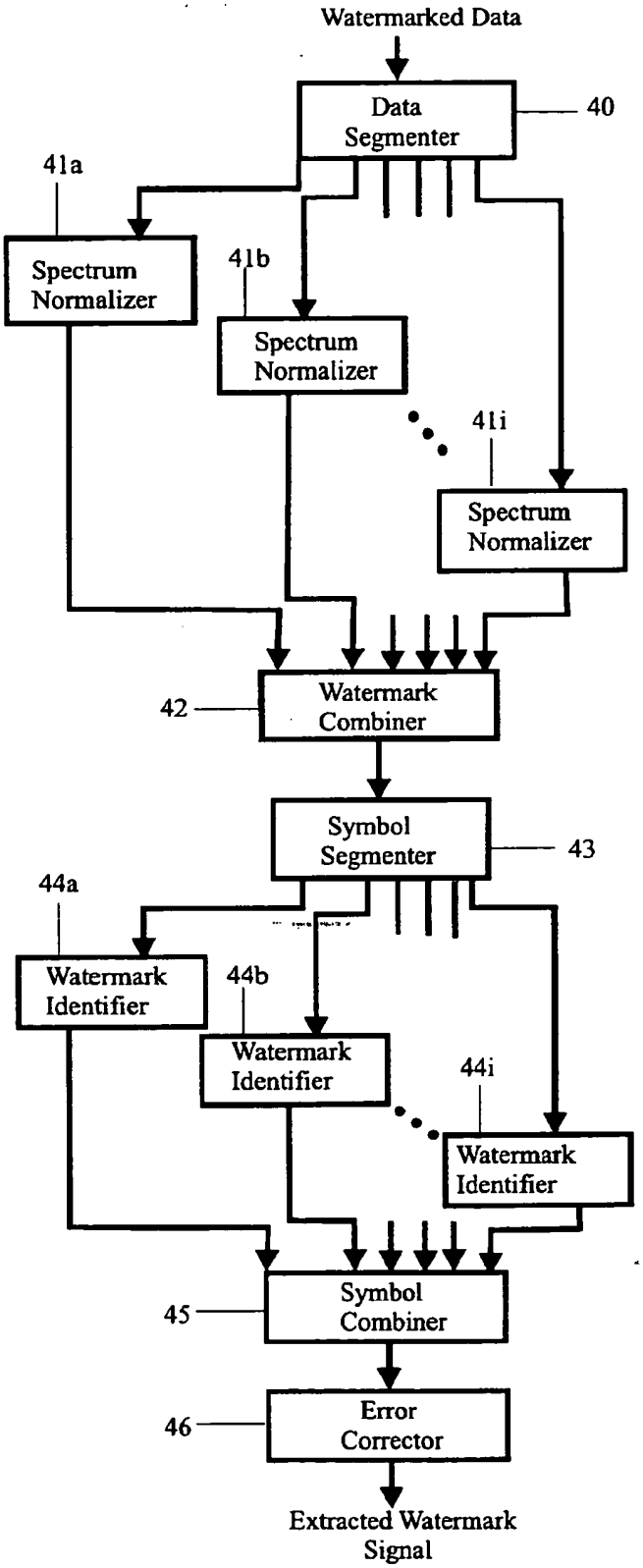

Watermarked Data
Data Segmenter
40
41a
Spectrum Normalizer
41b
Spectrum Normalizer
41i
Spectrum Normalizer
42
Watermark Combiner
Symbol Segmenter
43
44a
Watermark Identifier
44b
Watermark Identifier
44i
Watermark Identifier
45
Symbol Combiner
46
Error Corrector
Extracted Watermark Signal

Figure 4

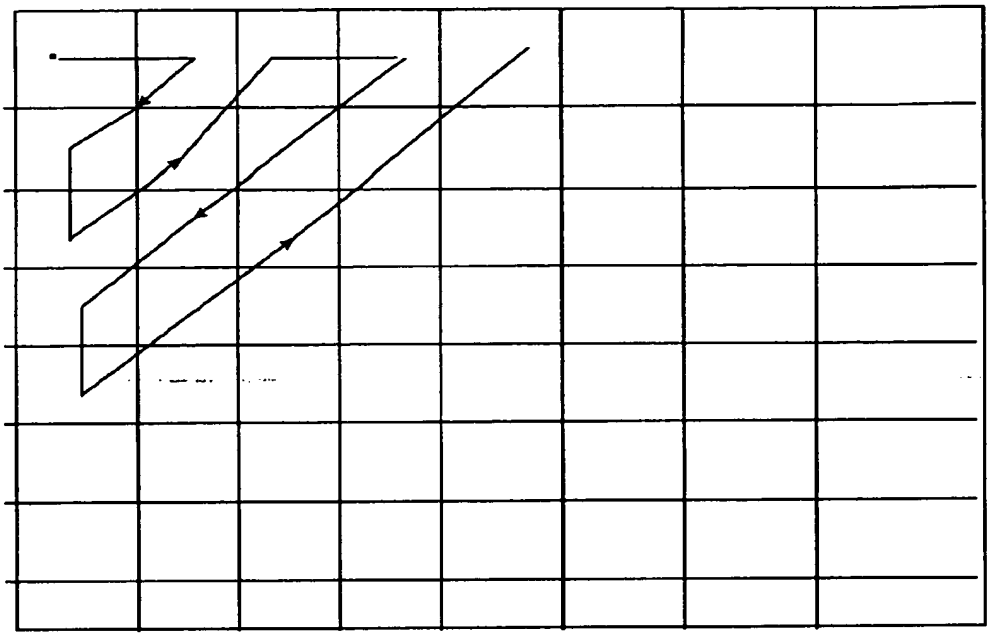


Figure 5

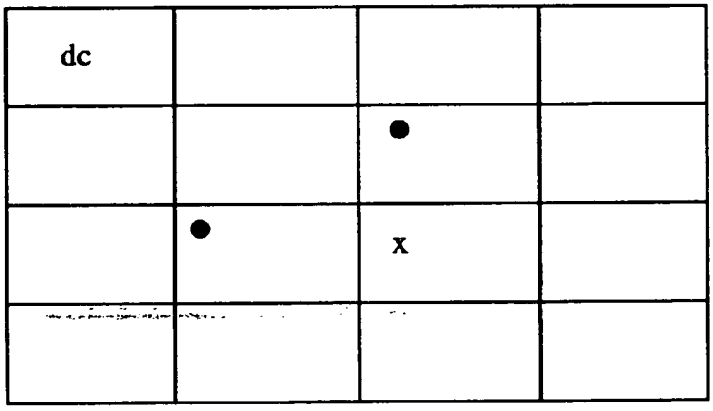


Figure 7

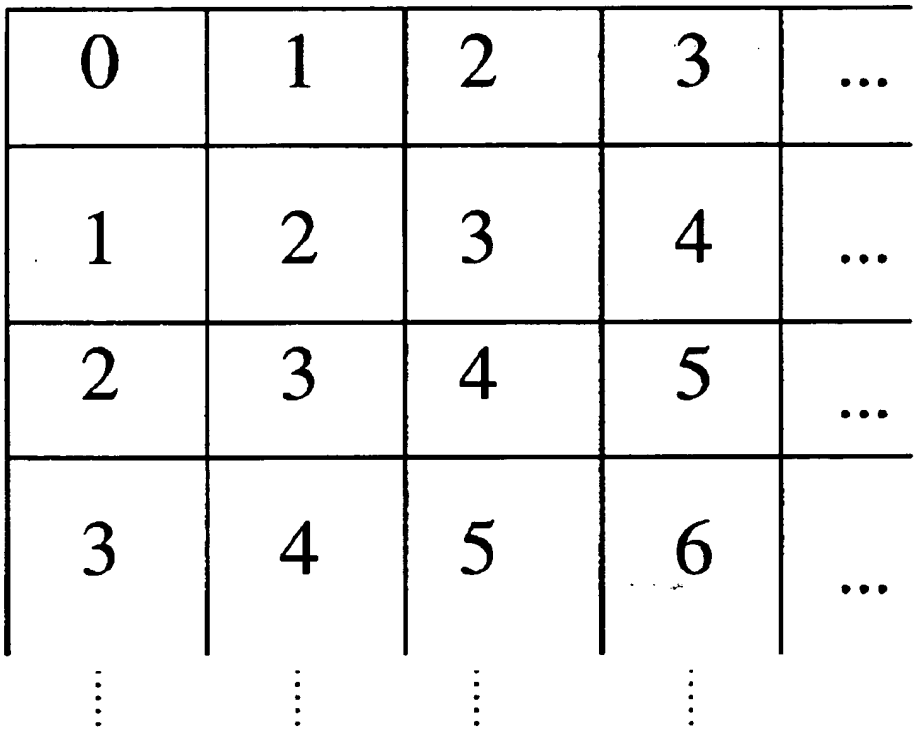


Figure 6

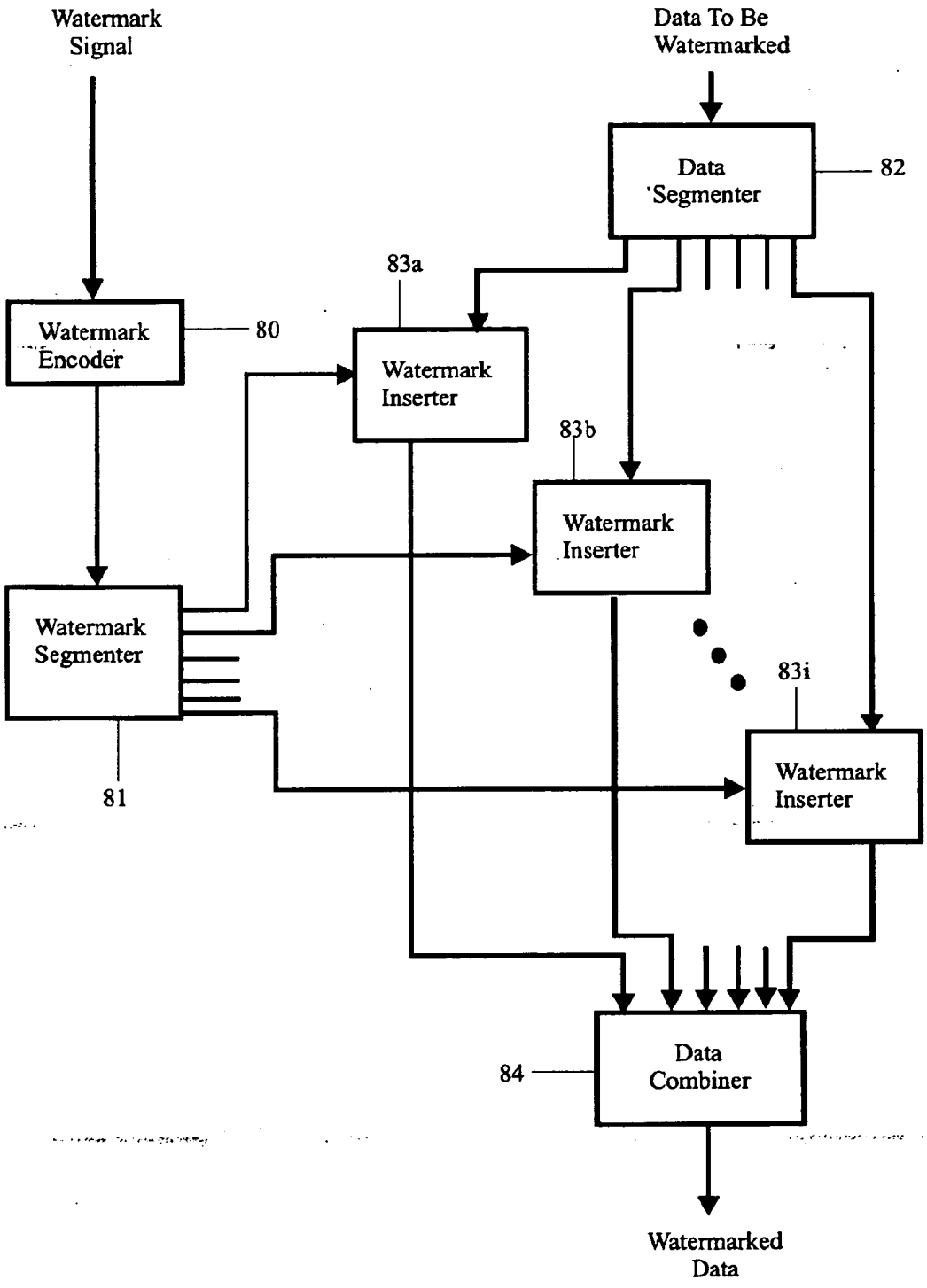

Watermark Signal
Data To Be Watermarked
Data Segmenter
82
83a
Watermark Encoder
80
Watermark Inserter
83b
Watermark Inserter
Watermark Segmenter
83i
Watermark Inserter
81
84
Data Combiner
Watermarked Data

Figure 8

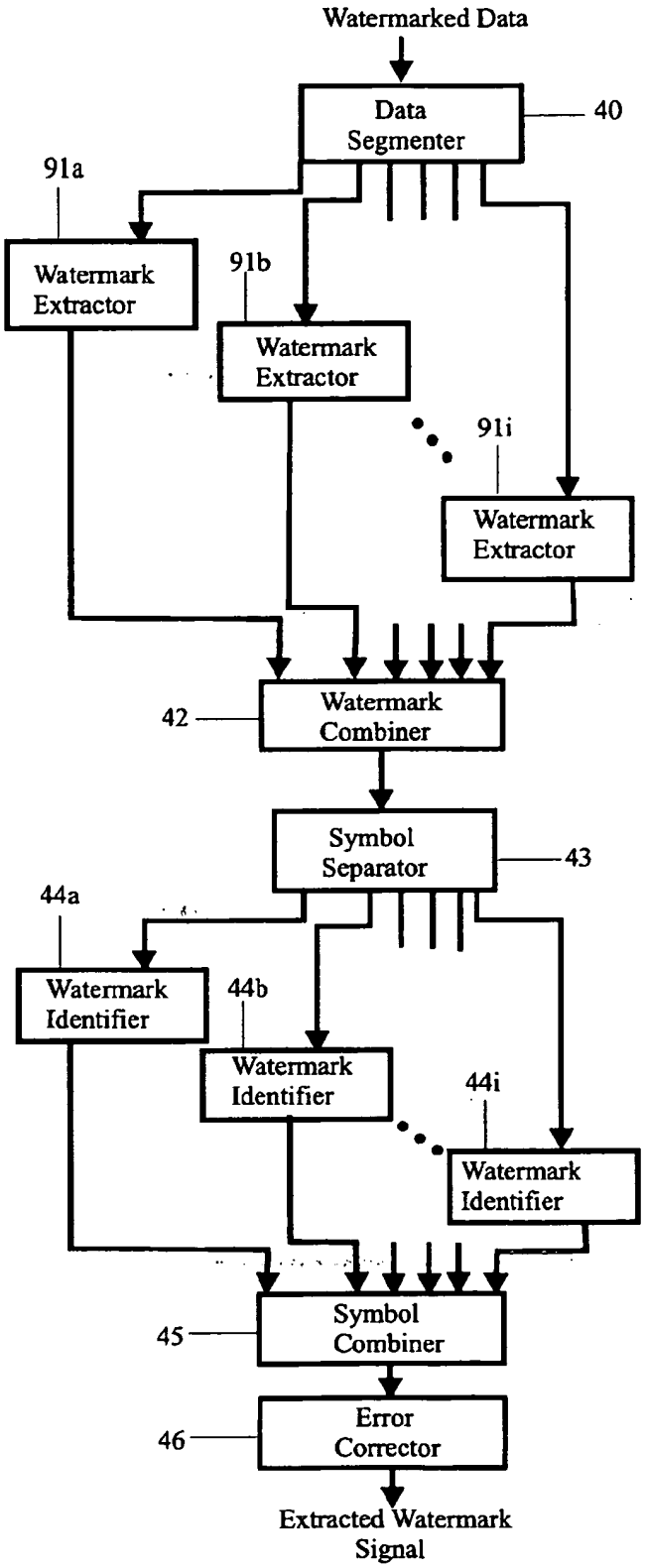

Watermarked Data
Data Segmenter
40
91a
Watermark Extractor
91b
Watermark Extractor
91i
Watermark Extractor
42
Watermark Combiner
Symbol Separator
43
44a
Watermark Identifier
44b
Watermark Identifier
44i
Watermark Identifier
45
Symbol Combiner
46
Error Corrector
Extracted Watermark Signal

Figure 9

DIGITAL WATERMARKING

FIELD OF INVENTION

The present invention relates to digital watermarking of data including image, video and multimedia data. Specifically, the invention relates to the insertion and extraction of embedded signals for purposes of watermarking, in which the insertion and extraction procedures are repeatedly applied to subregions of the data. When these subregions correspond to the 8×8 pixel blocks used for MPEG and JPEG compression and decompression, the watermarking procedure can be tightly coupled with these compression algorithms to achieve very significant savings in computation.

BACKGROUND OF THE INVENTION

The proliferation of digitized media such as image, video and multimedia is creating a need for a security system which facilitates the identification of the source of the material.

Content providers, i.e. owners of works in digital data form, have a need to embed signals into video/image/multimedia data which can subsequently be detected by software and/or hardware devices for purposes of authenticating copyright ownership, control and management.

For example, a coded signal might be inserted in data to indicate that the data should not be copied. The embedded signal should preserve the image fidelity, be robust to common signal transformations and resistant to tampering. In addition, consideration must be given to the data rate that can be provided by the system, though current requirements are relatively low—a few bits per frame.

In U.S. patent application Ser. No. 08/534,894, filed Sep. 28, 1995,. entitled "Secure Spread Spectrum Watermarking for Multimedia Data" now abandoned and assigned to the same assignee as the present invention, which is incorporated herein by reference, there was proposed a spread spectrum watermarking method which embedded a watermark signal into perceptually significant regions of an image for the purposes of identifying the content owner and/or possessor. A strength of this approach is that the watermark is very difficult to remove. In fact, this method only allows the watermark to be read if the original image or data is available for comparison. This is because the original spectrum of the watermark is shaped to that of the image through a non-linear multiplicative procedure and this spectral shaping must be removed prior to detection by matched filtering and the watermark is inserted into the N largest spectral coefficients, the ranking of which is not preserved after watermarking. Thus, this method does not allow software and hardware devices to directly read embedded signals.

In an article by Cox et al., entitled "Secured Spectrum Watermarking for Multimedia" available at http://www.neci.nj.com/tr/index.html (Technical Report No. 95-10) spread spectrum watermarking is described which embeds a pseudo-random noise sequence into the digital data for watermarking purposes.

The above prior art watermark extraction methodology requires the original image spectrum be subtracted from the watermark image spectrum. This restricts the use of the method when there is no original image or original image spectrum available. One application where this presents a significant difficulty is for third party device providers desiring to read embedded information for operation or denying operation of such a device.

In U.S. Pat. No. 5,319,735 by R. D. Preuss et al entitled "Embedded Signalling" digital information is encoded to produce a sequence of code symbols. The sequence of code symbols is embedded in an audio signal by generating a corresponding sequence of spread spectrum code signals representing the sequence of code symbols. The frequency components of the code signal being essentially confined to a preselected signaling band lying within the bandwidth of the audio signal and successive segments of the code signal corresponds to successive code symbols in the sequence. The audio signal is continuously frequency analyzed over a frequency band encompassing the signalling band and the code signal is dynamically filtered as a function of the analysis to provide a modified code signal with frequency component levels which are, at each time instant, essentially a preselected proportion of the levels of the audio signal frequency components in corresponding frequency ranges. The modified code signal and the audio signal are combined to provide a composite audio signal in which the digital information is embedded. This component audio signal is then recorded on a recording medium or is otherwise subjected to a transmission channel. Two key elements of this process are the spectral shaping and spectral equalization that occur at the insertion and extraction stages, respectively, thereby allowing the embedded signal to be extracted without access to the unwatermarked original data.

In U.S. patent application Ser. No. 08/708,331, filed Sep. 4, 1996,entitled "A Spread Spectrum Watermark for Embedded Signaling" by Cox; now U.S. Pat. No. 5,848,155 and incorporated herein by reference, there is described a method for extracting a watermark of embedded data from watermarked images or video without using an original or unwatermarked version of the data. This work can be viewed as an extension of the original work of Preuss et al from the audio domain to images and video.

This method of watermarking an image or image data for embedding signaling requires that the DCT (discrete cosine transform) and its inverse of the entire image be computed. There are fast algorithms for computing the DCT in N log N time, where N is the number of pixels in the image. However, for N=512×512, the computational requirement is still high, particularly if the encoding and extracting processes must occur at video rates, i.e. 30 frames per second. This method requires approximately 30 times the computation needed for MPEG-II decompression.

One possible way to achieve real-time video watermarking is to only watermark every N^{th} frame. However, content owners wish to protect each and every video frame. Moreover, if it is known which frames contain embedded signals, it is simple to remove those frames with no noticeable degradation in the video signal.

In U.S. patent application Ser. No. 08/715,953, filed Sep. 19, 1996,entitled "Watermarking of Image Data Using MPEG/SPEG Coefficients" by Cox, and incorporated herein by reference, there is described an alternative method, which is to insert the watermark into n×n blocks of the image (subimages) where n<<N. Then the computation cost is

$$\frac{N}{n} n \log n = N \log n.$$

For N=512×512=2^{18} and n=8×8=2^6, the asymptotic saving is only a factor of 3. However, empirically the cost of computing the DCT over the entire image may be significantly higher when cache, loop unfolding and other efficiency issues are considered. Thus, the practical difference

may approach a 30 fold savings. More importantly, if the block size is chosen to be 8×8, i.e. the same size as that used for MPEG image compression, then it is possible to tightly couple the watermark insertion and extraction procedures to those of the MPEG compression and decompression algorithms. Considerable computational saving can then be achieved since the most expenses computations relate to the calculation of the DCT and its inverse and these steps are already computed as part of the compression and decompression algorithm. The incremental cost of watermarking is then very small, typically less than 5% of the computational requirements associated with MPEG.

The present invention improves the reliability of the invention described in the 08/715,953 application, now pending by storing watermark information into subimages, and extracting watermark information from subimages, in a manner different from that described earlier.

SUMMARY OF THE INVENTION

The present invention improves the reliability of the prior systems by systematically varying the order in which watermark signal components are inserted into each subimage, by inserting only part of the watermark signal into each subimage, and, during watermark detection, by combining the watermark signals found in groups of subimages to reconstruct the original watermark signal before testing for correlation with any predefined watermarks.

For detection, a reverse transformation is applied to each subimage to reconstruct the watermark information that was stored in that subimage. The resulting signals are then averaged together to reconstruct the whole watermark, and to reduce noise. Finally, this reconstructed watermark is compared against a predefined set of watermark signals to determine which one was inserted into the image.

A principal object of the present invention is therefore, the provision of inserting a subset of a watermark into a subset of subregions of data to be watermarked.

Another object of the invention is the provision of a digital watermarking system in which a watermark is extracted by averaging the watermarked signal from subregions of watermarked data, and then correlating the resulting signal to determine the watermark.

A further object of the invention is the provision of a digital watermarking system in which the watermark is composed of two portions, a verification portion and a synchronization portion, in order to improve watermark extraction reliability

Further and still other objects of the invention will become more clearly apparent when the following description is read in conjunction with the accompanying drawing.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic block diagram of watermark insertion procedure;

FIG. 2 is a schematic block diagram of a watermark insertion procedure in accordance with the teachings of the present invention;

FIG. 3 is a schematic block diagram of a watermark extraction procedure;

FIG. 4 is a schematic block diagram of a watermark extraction procedure in accordance with the teachings of the present invention;

FIG. 5 is a graphic representation of a zigzag pattern useful for vectorizing subimages;

FIG. 6 is a graphic representation of rotation of PN sequences;

FIG. 7 is a graphical representation of an 8×8 block shown the spatial relation of averaged terms;

FIG. 8 is a schematic block diagram of a method for inserting watermarks in accordance with the present invention; and

FIG. 9 is a schematic block diagram of a method for extracting watermarks in accordance with the present invention.

DETAILED DESCRIPTION

Referring now to the figures, and FIGS. 1 through 4 in particular, there is shown schematic block diagrams of a general method for inserting and detecting watermarks in digital data, for instance images.

In the following description, reference may be made to image data or images. While the invention has applicability to image data and images, it will be understood that the teachings herein and the invention itself are equally applicable to video, image and multimedia data and the term "image" and "image data" will be understood to include these terms where applicable. As used herein, "watermark" will be understood to include embedded data, symbols, images, instructions or any other identifying information.

In the following description, reference is made to procedures described in U.S. patent application Ser. No. 08/534,894 for inserting and extracting or detecting a watermark in images as INSERT-ORIGINAL and EXTRACT-ORIGINAL, respectively. Reference is made to procedures described in U.S. patent application Ser. No. 08/708,331 filed Sep. 4, 1996,now U.S. Pat. No. 5,848,155 for inserting and extracting or detecting watermarks in images as INSERT-WHOLE and EXTRACT-WHOLE, respectively. And reference is made to procedures described in U.S. patent application Ser. No. 08/715,953 for inserting and extracting or detecting watermarks in images as INSERT-MPEG-A and EXTRACT-MPEG-A, respectively.

FIG. 1 shows a schematic block diagram of INSERT-WHOLE procedure for inserting watermarks into images. The watermark signal, in the form of a finite sequence of symbols chosen from an alphabet, is provided as an input to an error correction encoder **10** which transforms this sequence into another sequence that contains redundant information. The output of encoder **10** is provided to a PN-mapper **11**, which maps each symbol of the encoded watermark into a pre-specified pseudo-random noise (PN) code. The output of the PN-mapper **11** is provided to a spectral transformer **12**, which converts the pseudo-random noise sequence into the frequency domain. The conversion preferably is by discrete cosine transform (DCT), however, fast fourier transform, wavelet type decomposition and the like may also be used for frequency conversion. Concurrently, the data to be watermarked is provided to another spectral transformer **13**. The outputs of the two spectral transformers **12** and **13** are then provided as inputs to a spectral shaper **14**, which modifies the spectral properties of the pseudo-random noise codes from spectral transformer **12** to mask the watermark when added to the image data. The spectrally transformed data to be watermarked, from spectral transformer **13**, is also provided as an input to a delay **15**. The output of the spectral shaper **14** is then added to the output of delay **15** at a summer **16**. The summer output is subject to an inverse transform **17**. The result of the inverse transform is watermarked data.

INSERT-MPEG-A differs from INSERT-WHOLE by segmenting the data to be watermarked into multiple blocks,

such as 8×8 pixel subimages or subregions. Each block of data then has the watermark inserted according to the above described method. That is, for each 8×8 subimage or subregion, a pseudo-random number (PN) sequence is inserted into the DCT coefficients after suitable spectral shaping. The procedure is repeated for all such subimages or subregions. The size of the subimage or subregion is preferably 8×8, but it can be of other sizes, such as 2×2, 3×3, 4×4 or 16×16.

FIG. **2** shows a schematic block diagram of a watermark insertion procedure in accordance with teachings of the present invention. The watermark signal is processed into a noise spectrum signal by the error correction encoder **20**, the PN mapper **21**, and the spectral transformer **22**, in the same manner as described in conjunction with FIG. **1**. However, unlike INSERT-WHOLE or INSERT-MPEG-A, the watermark is then used as an input to a watermark segmenter **23**, which systematically separates the watermark into several subwatermarks. Any portion of the original watermark might appear redundantly in several of the resulting subwatermarks. Concurrently, the data to be watermarked is used as an input to data segmenter **24**, which segments the data into blocks or subregions, such as 8×8 subimages, as in INSERT-MPEG-A. Each of the subwatermarks output by the watermark segmenter **23** is then inserted into a data block by one of the watermark inserters **25***a*, **25***b*, etc. The procedure used by the watermark inserters **25***a*, **25***b*, etc., is the same procedure described connection with watermark inserter **18** in FIG. **1**. That is, each subwatermark is added into a spectrally transformed data block after spectral shaping, and the resulting data is then transformed back into the spatial domain. Finally, the watermarked data blocks from the watermark inserters **25***a*,**25***b*, etc., are assembled by data combiner **26** to produce watermarked data.

FIG. **3** shows a schematic block diagram of the EXTRACT-WHOLE procedure. The watermarked image, video or multimedia data is first used as input into a spectral normalizer **30** to undo any previously performed spectral shaping. If the data contains a watermark, then the output of the spectral normalizer **30** will resemble the spectral transformation of the PN coding of that watermark (the signal that was input to the spectral shaper **14** in FIG. **1**). The output of the spectral normalizer **30** is then used as an input to several correlators **31***a*, **31***b*, etc., which test the watermark with the PN codes used to represent the various symbols that the encoded watermark might contain (i.e. each correlator tests for one PN code that is used to encode a symbol by the PN mapper **11** of FIG. **1**). The outputs of the correlators **31***a*, **31***b*, etc., are used as inputs to a decision circuit **32**, which determines the most likely sequence of symbols. Finally, this sequence is corrected by an error corrector **33**, which performs the inverse coding that was performed by the error correction encoder **10** in FIG. **1**. The result is the extracted watermark.

In EXTRACT-MPEG-A, the data from which a watermark is to be extracted is first segmented into several blocks, such as 8×8 subimages, exactly as in INSERT-MPEG-A. The signal from each subimage is then normalized and used as input into a bank of correlators similar to the correlators **31***a*, **31***b*, etc. in FIG. **3**. The output from the correlators is then averaged with the outputs of corresponding correlators from other subimages, and the resulting average correlations are used as inputs into the decision circuit **32** for subsequent processing as described above.

FIG. **4** shows a schematic block diagram of a watermark extraction procedure in accordance with the teachings of the present invention. The watermarked data is first segmented into blocks by data segmenter **40**, which corresponds to the data segmenter **24** used during the insertion procedure in FIG. **2**. Each of the data blocks is provided to a respective spectrum normalizer **41***a*, **41***b*, etc. to produce a signal resembling the subwatermark that was inserted into the respective data block. These inserted subwatermark signals are then used as inputs into a watermark combiner **42**. In the combiner **42**, parts of the watermark that appear redundantly in several subwatermarks are averaged together to reduce noise. The output of the watermark combiner **42** is provided as the input to a symbol separator **43** which divides the watermark into parts, each of which corresponds to one symbol from the encoded watermark signal (the output of error correction encoder **20** in FIG. **2**).

These symbols from separator **43** are provided as inputs to respective watermark identifiers **44***a*, **44***b* etc. each of which includes of a bank of correlators and a decision circuit, as shown in FIG. **3**. The outputs of the watermark identifiers are symbols from the alphabet used in the original encoded watermark signal. The identified symbols are reassembled into a complete encoded watermark by the symbol combiner **45**. Finally, the resulting encoded watermark is decoded by the error corrector **46**.

The insertion and extraction procedures will now be described in more detail. In INSERT-ORIGINAL and EXTRACT-ORIGINAL, the object is to embed a single PN (pseudo random number) sequence into an image when the original image is available at the time of extraction. The information associated with the PN sequence is assumed to be stored in a database together with the original image and the spectral location of the embedded watermark. The locations of the watermarked components has to be recorded because the implementation approximated the N perceptually most significant regions of the watermark by the N largest coefficients. However, this ranking was not invariant to the watermarking process. The N largest coefficients may be different after inserting the watermark than before inserting the watermark.

In order to avoid this problem, the present invention places a watermark in predetermined locations of the spectrum, typically the first N coefficients. However, any predetermined locations could be used, though such locations should belong to the perceptually significant regions of the spectrum if the watermark is to survive common signals transformations such as compression, scaling, etc.

More generally, the information to be embedded is a sequence of m symbols drawn from an alphabet A (e.g. the binary digits or the ASCII symbols). This data is then supplemented with additional symbols for error detection and correction. Each symbol is then spread spectrum modulated, a process that maps each symbol into a unique PN sequence known as a chip. The number of bits per chip is preset - the longer the chip length, the higher the detected signal-to-noise ratio will be, but this is at the expense of signaling bandwidth.

The power spectrum of the PN sequence is white, i.e. flat, and is therefore shaped to match that of the "noise", i.e. the image/video/audio/or multimedia data into which the watermark is to be embedded. It is this spectral shaping that must be modified from the prior methods so that the extraction process no longer requires the original image. To do this, each coefficient of the watermarked spectrum is scaled by the local average of the power in the image spectral coefficient rather than the coefficient itself, i.e.

$$f_i' = f_i + \alpha \mathrm{avg}(|f_i|) W_i \quad (1)$$

The averaging is the averaging of the absolute coefficient values and not the coefficient values themselves. This is effectively estimating the average power present at each frequency. Other averaging procedures are possible, for example, averaging over several frames or average of local neighborhoods of 8×8 blocks.

This average may be obtained in several ways. It may be a local average over a two dimensional region. Alternatively, the two dimensional spectrum may be sampled to form a one dimensional vector and a one dimensional local average may be performed. One dimensional vectorization of the two dimensional 8×8 DCT coefficients is already performed as part of MPEG II. The average may be a simple box or weighted average over the neighborhood.

For video data, temporal averaging of the spectral coefficients over several frames can also be applied. However, since several frames are needed for averaging at the spectral normalization stage of the extractor, the protection of individual video frames taken in isolation may not be possible. For this reason, the present invention treats video as a very large collection of still images. In this way, even individual video frames are copy protected.

In order to extract the watermark, it is necessary to perform the spectral normalization, in which the previously performed spectral shaping procedure is inverted. In the present invention, the original unwatermarked signal is not available. Thus, the average power of the frequency coefficients, avg($|f_i|$), is approximated by the average of the watermarked signal, i.e. avg($|f_i'|$)

$$\text{avg}(|f_i|) \approx \text{avg}(|f_i'|) \tag{2}$$

This is approximately true since $\alpha \text{avg}(|f_i|)W_i << f_i$, where W_i is the watermark component, and α is a constant typically in the range between 0.1 and 0.01.

The normalization stage then divides each coefficient (f_i') in the received signal by the local average avg ($|f_i'|$) in the neighborhood.

That is,

$$\frac{f_i'}{\text{avg}(|f_i'|)} = \frac{f_i + \alpha \text{avg}(|f_i|)W_i}{\text{avg}(|f_i|)} \approx \frac{f_i'}{\text{avg}(|f_i'|)} + \alpha W_i \tag{3}$$

The first term, on the right hand side (RHS) of Equation (3),

$$\frac{f_i}{\text{avg}(|f_i'|)},$$

is considered a noise term. This term was not present in the system described in U.S. patent application Ser. No. 08/534,894, because access to the unwatermarked coefficients allowed this term to be removed. The second term αW_i is the original watermark signal which can now be detected using conventional correlation.

If the watermark is extracted from any single 8×8 block, the detector reliability is very low. If, however, the watermarks extracted from each 8×8 block are first added together and the averaged watermark is then applied to the correlator, then a very strong and unambiguous response is obtained. This differs from the method described in U.S. patent application Ser. No. 08/715,953 in which correlation occurred within each block and the output from each correlator was averaged together. The present invention was found to improve the detection response and significantly reduced the computation requirement associated with each block.

In practicing the present invention preferably there is a unique PN sequence for each symbol in the alphabet. The method is relatively robust to clipping since the detector output reduces linearly with the quantity of 8×8 subimage blocks in the image. For DVD (digital video disk) embedded signaling for APS (analog protection system) and CGMS (copy generation management system), there would be a total of 8 or 16 PN sequences.

The number of 8×8 blocks in a 512×512 image is 4096, suggesting that significantly more than one of 16 symbols can be embedded in an image or video frame. Assume, for example, that it is desired to embed 1 out of 128 symbols in an image. It is necessary to perform 128 parallel correlations. This is computationally tractable but hardware implementations of each correlation become more complex. An alternative method is to only use two binary symbols. It may be preferable to associate more than one PN sequence with each of the two binary symbols or bits in order to increase the difficulty of intentionally removing the watermark. In this case, there are only two correlators and a binary string may be embedded into the image. The raw bit error rate will be very high due by the low detector output. However, this can be reduced to acceptable levels by using error correcting codes, such as Reed-Solomon (RS). RS codes are robust to burst error which may occur because of clipping of the image. Other error correcting codes may also be used.

When using this method, it is necessary for the receiver to know the start location of the encoded block. The start location may not be obvious, particularly when the image has been subjected to clipping. However, convention synchronizing methods can be used; such as preceding each block with a special or unique symbol or string of symbols.

To insert a watermark, each 8×8 block is treated as an individual subimage or subregion. The DCT of the subimage is then computed and the two dimensional DCT is vectorized in the zigzag pattern shown in FIG. 5, although other patterns are also possible. These two stages constitute most of the calculations but are part of the MPEG encoding process. Next, a PN noise sequence $\{w_1 \ldots w_n\}$ is inserted into the DCT coefficients using Equation 1 as before. The length of the PN sequence cannot exceed 64 (in an 8×8 block) and is typically much shorter, in the range of 11 to 25. If only a single code is to be inserted into the image, then the same PN sequence is inserted into each of the 720 x 480/64 –5400 blocks. However, a variation may be performed at this point in the procedure. Within each row of blocks, the PN sequence is cyclically rotated by one frequency coefficient prior to insertion in the subsequent block. Similarly, the PN sequence is cyclically rotated by one frequency coefficient at the start of each new row. FIG. 6 illustrates an order of rotations.

The purpose of these rotations or shifts is to improve the response of the watermark extraction stage. Earlier experiments revealed that certain DCT coefficients were more difficult to estimate than others. The location at these coefficients varied from image to image. However, within an image, the coefficient could be consistently poor. Consequently, without shifting, one or more of the estimated watermark coefficients could be significantly degraded relative to the other watermark coefficients, thereby reducing the detector performance. Conversely, shifting significantly reduces the effect a poor DCT coefficients has on a single watermark coefficient and the detector performance is markedly improved. Note that any cyclic pattern can be used.

Further modifications are useful once rotation of the watermark has been introduced. First, the length of the watermark may now be significantly greater than 64.Then, for each block only a small subset of the watermark (say five) coefficients is inserted into the first five DCT coefficients (excluding the d.c. term). Because of the rotation, a different subset of the watermark is inserted into neighboring 8×8 blocks. Finally, having completed the watermark insertion, the MPEG encoder is able to proceed with the subsequent stages of compression.

Note that the watermark may also be inserted after the MPEG quantization stage to reduce distortion of the watermark. MPEG-2 performs a convenient one dimension vectorization called "zigzagging", which allows a simple 3×1 box average to be performed on the coefficients (excluding the d.c. term).

In practice, performance was improved if the averaging is performed using the 2 four-connected coefficients closest to the d.c. term, as illustrated in FIG. 7, i.e. the two coefficients above and to the left.

Watermark detection begins by first extracting the PN noise sequence from each 8×8 block using Equation 1.For each block, the PN sequence is then cyclically shifted in the opposite direction by one frequency coefficient, and the average over all the blocks is then computed. In practice, this process can be computed incrementally and does not require temporary storage of all the extracted watermarks. A weighted averaging can also be applied, where the weights are determined based on their susceptibility to common signal transformations such as low pass filtering. Finally, the average watermark is compared with the original PN sequence via correlation. The reason for shifting the watermark in the column direction may now be apparent. If the image is clipped on an arbitrary block boundary, then the computed average watermark will simply be rotated by an amount that is a function of the relative location of the clipped portion of the image. Correlation can then be performed on all permutations (typically 11 to 25) of the watermark. The output from the correlator with the maximum value is then used for decision purposes. The extraction stage is depicted in FIG. 4. Taking the maximum correlator output over all rotations of the watermark can cause the decision circuitry to be noisy. To improve this, the watermark is broken into two pieces; a synchronization portion is of length K and a verification portion is N-K. Then, when the watermark is extracted as before, correlation is first performed only on all rotations of the synchronization portion of this watermark. The maximum correlation output is noted, then the verification portion of the watermark is rotated by the corresponding amount and a second correlation is performed on the verification portions of the watermarks. This process significantly improves the overall reliability of the system. In the course of experimentation, it was noticed that some watermarks performed better than others on the same imagery. This was caused by variation in the correlation statistics between the synchronization and verification portions of the watermark. Ideally, the two portions should have very low correlations. However, in several cases where watermarks performed poorly, it was traced to unexpected correlations between the two portions.

The present invention provides a modification to digital watermarking methods in which the original data is required for watermark extraction thereby enabling watermarking extraction in the absence of an unwatermarked or original data. The present invention preferably uses MPEG/JPEG coefficients. An image is divided into typically 8×8 block subimages or subregions and each subimage is processed and the results are combined to derive the extracted watermark. The result is extraction of the watermark with very high confidence.

While the above invention describes improvements to the prior-art INSERT-WHOLE, INSERT-MPEG-A, EXTRACT-WHOLE, and EXTRACT-MPEG-A algorithms, it should be apparent to anyone skilled in the art that the same improvements may be applied to any algorithm for inserting and extracting watermarks in image data. This more general view of the present invention is shown in FIGS. **8** and **9**.

FIG. **8** shows a schematic block diagram of the general method for inserting watermarks. This general method makes use of a non-block-based watermark insertion algorithm, which shall be referred to hereafter as the "base insertion algorithm". The watermark encoder **80** converts the watermark into a form appropriate for the base insertion algorithm. If the base insertion algorithm is that shown in FIG. **1**, for example, then the watermark encoder **80** corresponds to the watermark encoder **19**, which comprises the error correction encoder **10**, the PN mapper **11**, and the spectral transformer **12**. However, if a different base insertion algorithm is to be used, then the watermark encoder **80** may perform a different transformation of the watermark. The encoded watermark signal from watermark encoder **80** is provided as an input to watermark segmenter **81**, which divides the watermark into a set of subwatermarks. Any portion of the original watermark might appear redundantly in several of the resulting subwatermarks. The data to be watermarked is provided as an input to data segmenter **82**, which divides the data into subregions. Each subwatermark is inserted into a respective data subregion by a watermark inserter **83***a*, **83***b*, etc. The watermark inserters implement the base insertion algorithm, so, if the base insertion algorithm is that shown in FIG. **1**, then each watermark inserter **83***a*, **83***b*, etc., corresponds to the watermark inserter **18**, which comprises a spectral transformer **13**, a spectral shaper **14**, a delay **15**, a summer **16**, and an inverse transform **17**. However, if a different base insertion algorithm is to be used, then the watermark inserters **83***a*, **83***b*, etc., may employ a different method of inserting subwatermarks into the subregions of the data to be watermarked. The outputs from the watermark inserters are assembled in data combiner **84** to provide watermarked data.

FIG. **9** shows a schematic block diagram of the corresponding general extraction algorithm. The algorithm makes use of a "base extraction" algorithm that corresponds to the base insertion algorithm used in inserting the watermark into the data to be watermarked (FIG. **8**). The algorithm in FIG. **9** is substantially the same as the algorithm shown in FIG. **4**, except that, in the general case, the spectrum normalizers **41***a*, etc. are replaced by watermark extractors **91***a*, etc., which implement the base extraction algorithm. That is, if the base insertion algorithm used was the algorithm shown in FIG. **1**, then the watermark extractors **91***a*, etc., in FIG. **9** will be the spectrum normalizers **41***a*, etc. in FIG. **4**.

While there has been described and illustrated a system for inserting a watermark into and extracting a watermark from watermarked data without using an unwatermarked version of the data, it will be apparent to those skilled in the art that variations and modifications are possible without deviating from the broad principles and teachings of the present invention which shall be limited solely by the scope of the claims appended hereto.

What is claimed is:

1. A method for inserting a watermark signal into data to be watermarked comprising the steps of:

dividing data to be watermarked into a plurality of subregions;

computing frequency coefficients of the data to be watermarked in each subregion;

spread spectrum modulating a watermark signal to be inserted by mapping the watermark signal into a PN (pseudo-random noise) sequence;

spectral shaping the PN sequence as a function of the average power in each frequency coefficient of the data; and

inserting each spectral shaped PN sequence into predetermined coefficients in the data in each subregion.

2. A method for inserting a watermark signal into data to be watermarked as set forth in claim **1**, where said inserting is performed after the data undergoes MPEG quantization processing.

3. A method for inserting a watermark signal into data to be watermarked as set forth in claim **1**, where said frequency coefficients are DCT (discrete cosine transform) coefficients.

4. A method for inserting a watermark signal into data to be watermarked as set forth in claim **3**, where each subregion is a 8×8 block of pixels.

5. A method for inserting a watermark signal into data to be watermarked as set forth in claim **4**, where said inserting is performed after the data undergoes MPEG quantization processing.

6. A method for inserting a watermark signal into data to be watermarked as set forth in claim **1**, where each subregion is a 8×8 block of pixels.

7. A method for inserting a watermark signal into data to be watermarked as set forth in claim **6**, where said inserting is performed after the data undergoes MPEG quantization processing.

8. A method for inserting a watermark signal into data to be watermarked as set forth in claim **6**, where the frequency coefficients of the watermark signal are rotated prior to inserting of each spectral shaped PN sequence into the subregion.

9. A method for inserting a watermark signal into data to be watermarked as set forth in claim **8**, where said inserting is performed after the data undergoes MPEG quantization processing.

10. A method for inserting a watermark signal into data to be watermarked as set forth in claim **8**, where only a subset of the watermark signal frequency coefficients is inserted into any one subregion.

11. A method for inserting a watermark signal into data to be watermarked as set forth in claim **10**, where the watermark signal comprises a synchronization portion and a verification portion.

12. A method for inserting a watermark signal into data to be watermarked as set forth in claim **11**, where said inserting is performed after the data undergoes MPEG quantization processing.

13. A method for inserting a watermark signal into data to be watermarked as set forth in claim **11**, where the synchronization portion and the verification portion have very little correlation between each other.

14. A method for inserting a watermark signal into data to be watermarked as set forth in claim **1**, where the spectral shaping as a function of the average power is typically 3×1 window of the coefficient obtained from the one-dimensional vectorization by zigzagging of two-dimension frequency coefficients.

15. A method for inserting a watermark signal into data to be watermarked as set forth in claim **1**, where the spectral shaping is a function of the average power based on the two four-connected frequency coefficients closest to the DC term.

16. A method of extracting a watermark from watermarked data comprising the steps of:

receiving subregions of watermarked data;

spectrum normalizing the watermarked data as a function of the average power in each frequency coefficient of the watermarked data in each subregion to generate respective normalized signals;

combining the respective normalized signals from each subregion to generate a single watermark;

correlating the single watermark with predetermined PN (pseudo-random noise) sequences corresponding to predetermined symbols to provide correlated signals for each predetermined PN sequence in each subregion;

deciding which correlated signal is most likely a current symbol; and

extracting a sequence of most likely current symbols corresponding to the watermark.

17. A method of extracting a watermark from watermarked data as set forth in claim **16**, where the subregions are 8×8 blocks used for MPEG encoding and decoding.

18. A method of extracting a watermark from watermarked data as set forth in claim **17**, where said combining the normalized signals from each subregion to generate a single watermark, including removing the relative rotation of the watermark between blocks.

19. A method of extracting a watermark from watermarked data as set forth in claim **18**, further comprising subsequently reconstructing the watermark from partial watermarks inserted into each block.

20. A method of extracting a watermark from watermarked data as set forth in claim **19**, further comprising weighting the watermark coefficients based on their location within the frequency spectrum, where the weighting is a function of the susceptibility of each frequency coefficient to common signal transformations.

21. A method of extracting a watermark from watermarked data as set forth in claim **16**, further comprising correlating with all rotational shifts of the extracted watermark and selecting the maximum value.

22. A method of extracting a watermark from watermarked data as set forth in claim **16**, further comprising correlating with all rotational shifts of a synchronization portion of a watermark to determine a maximum value and subsequently rotating a verification portion of the watermark by the same amount as the synchronization portion is rotated to obtain the maximum value prior to correlating between the verification portion and predetermined PN sequences.

23. A method of extracting a watermark from watermarked data comprising the steps of:

receiving subregions of watermarked data;

spectrum normalizing the watermarked data as a function of the average power in each frequency coefficient of the watermarked data in each subregion to generate respective normalized signals;

correlating the respective normalized signals with predetermined PN sequences corresponding to predetermined symbols to provide correlated signals for each predetermined PN sequence in each subregion;

deciding which correlated signal is most likely a current symbol in each subregion for providing an extracted symbol stream;

error correcting the extracted symbol stream; and

extracting a sequence of most likely current symbols corresponding to the watermark.

24. A method of extracting a watermark from watermarked data as set forth in claim **23**, where said error correction is Reed Solomon error correction.

25. A method for inserting a watermark signal into data to be watermarked comprising the steps of:

dividing data to be watermarked into a plurality of subregions;

dividing a watermark signal into a plurality of subwatermarks where portions of the watermark are contained in more than one subwatermark; and

inserting said plurality of subwatermarks into said plurality of subregions.

26. A method for inserting a watermark signal into data to be watermarked as set forth in claim 25, where each subwatermark is inserted into a respective subregion, so that each subregion contains at least one subwatermark.

27. A method for extracting a watermark signal from watermarked data comprising the steps of:

receiving a plurality of subregions of watermark data;

extracting a subwatermark from each subregion of said plurality of subregions; and

combining and averaging the subwatermarks extracted from all the subregions to obtain a signal commensurate with the watermark signal.

28. A method for extracting a watermark signal from watermarked data as set forth in claim 27, further comprising the steps of:

dividing the signal commensurate with the watermark signal into a plurality of symbol signals;

correlating each symbol signal with a set of predefined signals;

determining which predefined signal best corresponds to each symbol signal; and

combining the best corresponding predetermined signals to generate the watermark signal.

* * * * *

US005930369A

United States Patent [19]

Cox et al.

[11] Patent Number: 5,930,369

[45] Date of Patent: Jul. 27, 1999

[54] SECURE SPREAD SPECTRUM WATERMARKING FOR MULTIMEDIA DATA

[75] Inventors: Ingemar J. Cox, Lawrenceville; Joseph J. Kilian, Princeton Junction; Talal G. Shamoon, Princeton, all of N.J.

[73] Assignee: NEC Research Institute, Inc., Princeton, N.J.

[21] Appl. No.: 08/926,720

[22] Filed: Sep. 10, 1997

Related U.S. Application Data

[63] Continuation of application No. 08/534,894, Sep. 28, 1995, abandoned.

[51] Int. Cl.[6] G09C 5/00; H04L 9/00

[52] U.S. Cl. 380/54; 380/3; 380/4; 380/23; 380/55; 283/73; 283/113; 283/17

[58] Field of Search 380/3, 4, 9, 23, 380/54, 59, 51, 55; 283/73, 113, 17, 901

[56] References Cited

U.S. PATENT DOCUMENTS

4,939,515	7/1990	Adelson	341/51
5,319,735	6/1994	Preuss et al.	395/2.14
5,488,664	1/1996	Shamir	380/54
5,530,751	6/1996	Morris	380/4
5,530,759	6/1996	Braudaway	380/54
5,568,570	10/1996	Rabbani	382/238
5,646,997	7/1997	Barton	380/23
5,659,726	8/1997	Sandford, II et al.	395/612
5,664,018	9/1997	Leighton	380/54
5,734,752	3/1998	Knox	380/54 X

FOREIGN PATENT DOCUMENTS

0690595	1/1995	European Pat. Off. .	
2196167	4/1998	United Kingdom .	
8908915	9/1989	WIPO	G11B 20/10
9514289	5/1995	WIPO	G06K 19/14
9520291	7/1995	WIPO .	
9621290	7/1996	WIPO	H04H 1/00
9625005	8/1996	WIPO	H04H 7/08
9627259	9/1996	WIPO .	

OTHER PUBLICATIONS

I. Cox et al, "Secure Spread Spectrum Watermarking for Images, Audio and Video", in IEEE Int. Conference On Image Processing, vol. 3, pp. 243–246, 1996.

I. Cox et al, "A Secure, Robust Watermark for Multimedia", in Information Hiding: First Int. Workshop Proc., R. Anderson, ed., vol. 1174 of Lecture notes in Computer Science, pp. 185–206, Springer–Verlag 1996 IEEE Int. Conf. On Image Processing, 1996.

J. Brassil et al, "Watermarking document images with bounding box expansion", in Information Hiding, R. Anderson, ed., vol. 1174, of Lecture Notes in Computer Science, pp. 227–235, Springer–Verlag, 1996.

J.R. Smith et al, "Modulation and information hiding in images", in information Hiding: First Int. Workshop Proc., R. Anderson, ed., vol. 1174 of Lecture Notes in Computer science, pp. 207–226, Springer–Verlag 1996.

R.L. Rivest et al, "A method for obtaining digital signatures and public–key cryptosystems", Communications Of the ACM, vol. 21, No. 2, Feb. 1978, pp. 120–126.

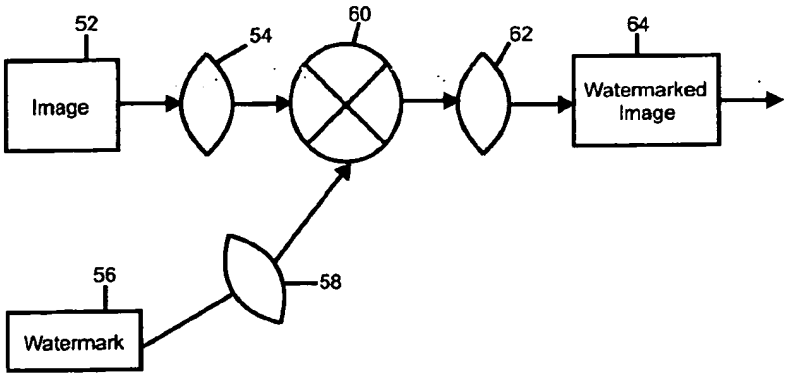
(List continued on next page.)

Primary Examiner—Bernarr E. Gregory
Attorney, Agent, or Firm—Philip J. Feig

[57] ABSTRACT

Digital watermarking of audio, image, video or multimedia data is achieved by inserting the watermark into the perceptually significant components of a decomposition of the data in a manner so as to be visually imperceptible. In a preferred method, a frequency spectral image of the data, preferably a Fourier transform of the data, is obtained. A watermark is inserted into perceptually significant components of the frequency spectral image. The resultant watermarked spectral image is subjected to an inverse transform to produce watermarked data. The watermark is extracted from watermarked data by first comparing the watermarked data with the original data to obtain an extracted watermark. Then, the original watermark, original data and the extracted watermark are compared to generate a watermark which is analyzed for authenticity of the watermark.

46 Claims, 6 Drawing Sheets



OTHER PUBLICATIONS

E. Franz et al, "Computer–based steganography: how it works and why therefore any restrictions on Crytography are nonsense, at best" in Information Hiding, First Int. workshop, Cambridge, UK. Jun. 1996.

Ingemar J. Cox et al., "A Review of watermarking and the importance of perceptual modeling" Proc. of IE '97, vol. 3016, Feb. 9–14, 1997.

Boland et al., "Watermarking Digital Images for Copyright Protection", Image Processing and Its Applications, Jul. 4–6 1995, Conference Publication No. 410.

Kahn et al., "Information Hiding—An Annotated Bibliography",Oct. 17, 1995.

R.G. Van Schyndel et al, "A digital watermark," in Intl. Conf. On Image Processing, vol. 2, pp. 86–90, 1994.

G. Caronni, "Assuring Ownership Rights for Digital Images," in Proc. Reliable IT Systems, VIS '95, 1995.

J. Brassil et al, "Electronic Marking and Identification Techniques to Discourage Document Copying," in Proc. Infocom '94, pp. 1278–1287, 1994.

K. Tanaka et al, "Embedding Secret Information into a Dithered Multi–Level Image," in IEEE Military Comm. Conf., pp. 216–220, 1990.

K. Mitsiu et al, "Video–Steganography: How to Secretly Embeded a Signature in a Picture," in IMA Intelluctual Property Project Proc., vol. 1, pp. 187–206, 1994.

Macq and Quisquater, "Cryptology for Digital TV Broadcasting," in Proc. of the IEEE, vol. 83, No. 6, pp. 944–957, 1995.

W. Bender et al, "Techniques for data hiding," in Proc. of SPIE, vol. 2420, No. 40. Jul. 1995.

Koch, Rindfery and Zhao, "Copyright Protection for Multimedia Data," in Proc. of the Int'l Conf. on Digital Media and Electronic Publishing (Leeds, UK, Dec. 6–8, 1994).

Kock and Zhao, "Towards Robust and Hidden Image Copyright Labeling," in Proc. of 1995 IEEE Workshop on Nonlinear Signal and Image Processing (Neos Marmaras, Halkidiki, Greece, Jun. 20–22, 1995.

Zhao and Koch, " Embedding Robust Labels Into Images For Copyright Protection,"in Proc. Int. Congr. on IPR for Specialized Information, Knowledge and New Technologies (Vienna, Austria), Aug. 21–25, 1995.

"Digital Copyright: Who Owns What?" NewMedia, Sep. 1995, pp. 38–43.

"Publish and Be Robbed?" New Scientist, Feb 18, 1995, pp. 32–37.

Kohno et al, "Spread Spectrum Access Methods for Wireless IEEE Communications," in Communications Magazine, Jan. 1995, pp. 58–67, 116.

Campana and Quinn, "Spread spectrum communications," in IEEE Potentials. Apr. 1993, pp. 13–16.

Mowbray and Grant, "Wideband coding for uncoordinated multiple access communications," in Electronics & Communication Engineering Journal, Dec. 1992, pp. 351–361.

Digimarc Overview & "Wired" Magazine article (Jul. 1995 issue)—Jun. 1995.

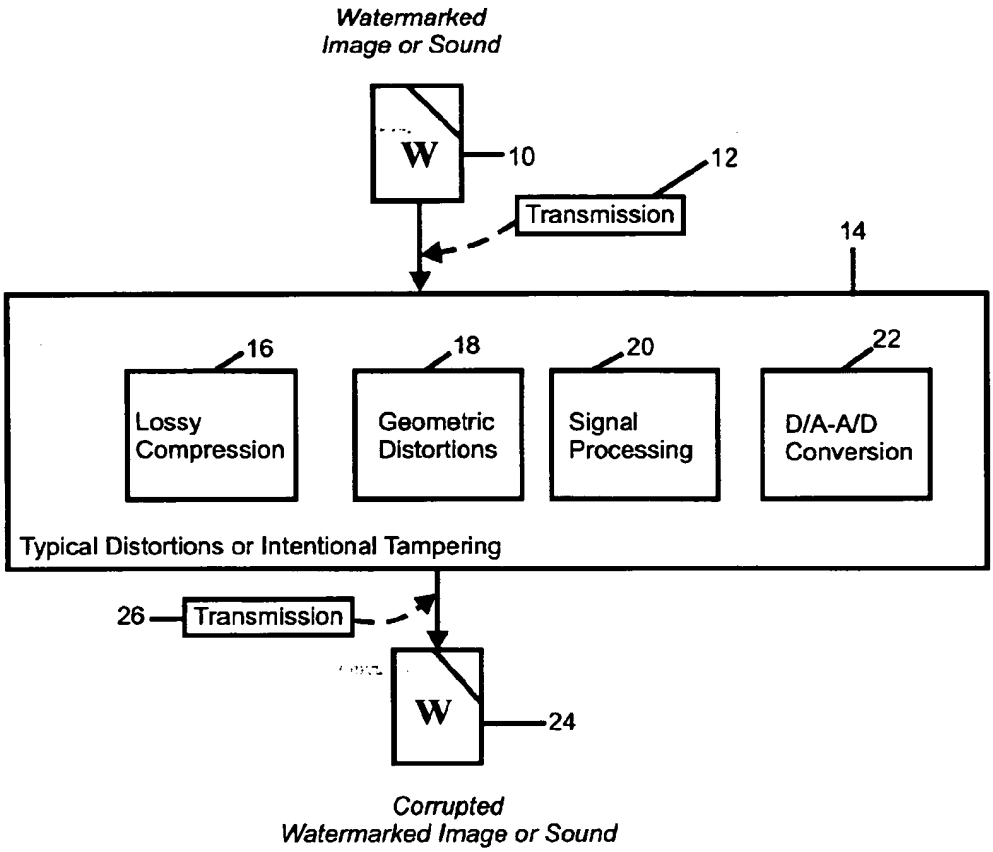

Watermarked
Image or Sound
W
10
12
Transmission
14
16
Lossy
Compression
18
Geometric
Distortions
20
Signal
Processing
22
D/A-A/D
Conversion
Typical Distortions or Intentional Tampering
26
Transmission
W
24
Corrupted
Watermarked Image or Sound

FIGURE 1

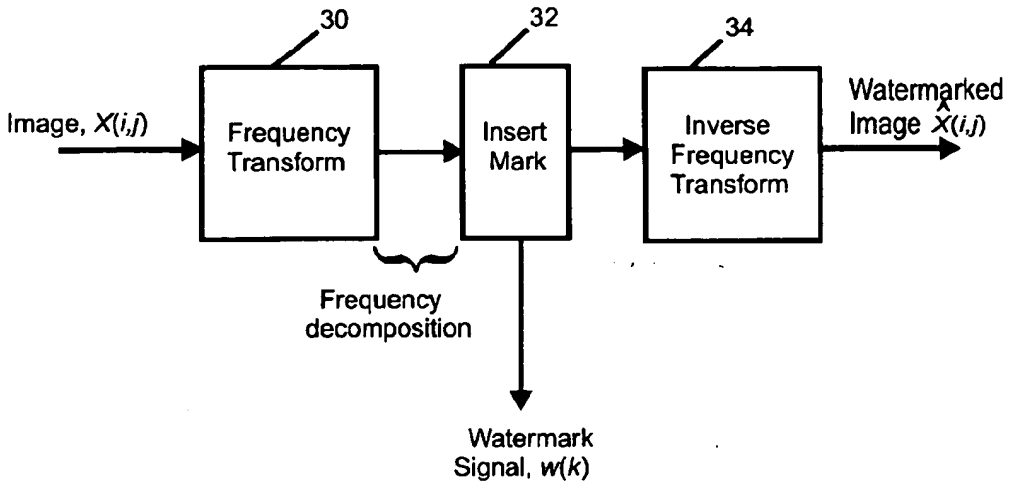


FIGURE 2

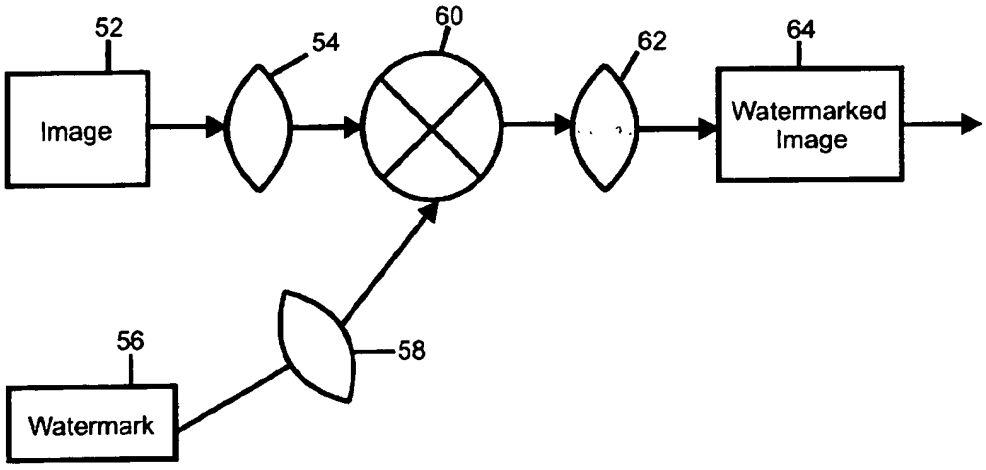


FIGURE 7

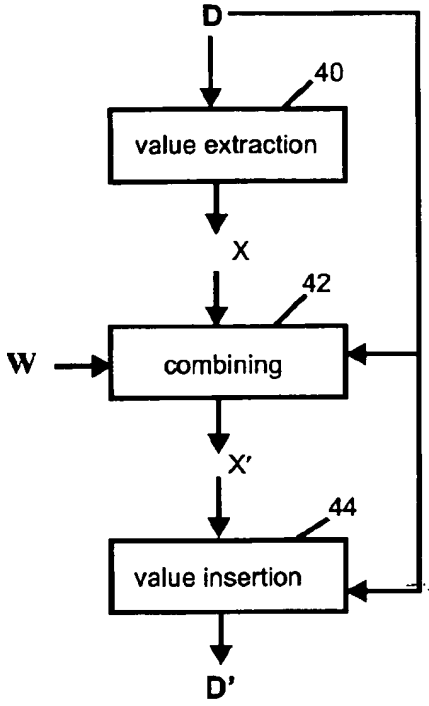


FIGURE 3a

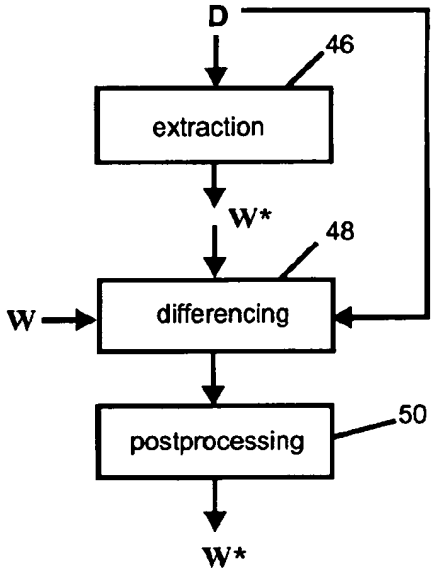


FIGURE 3b

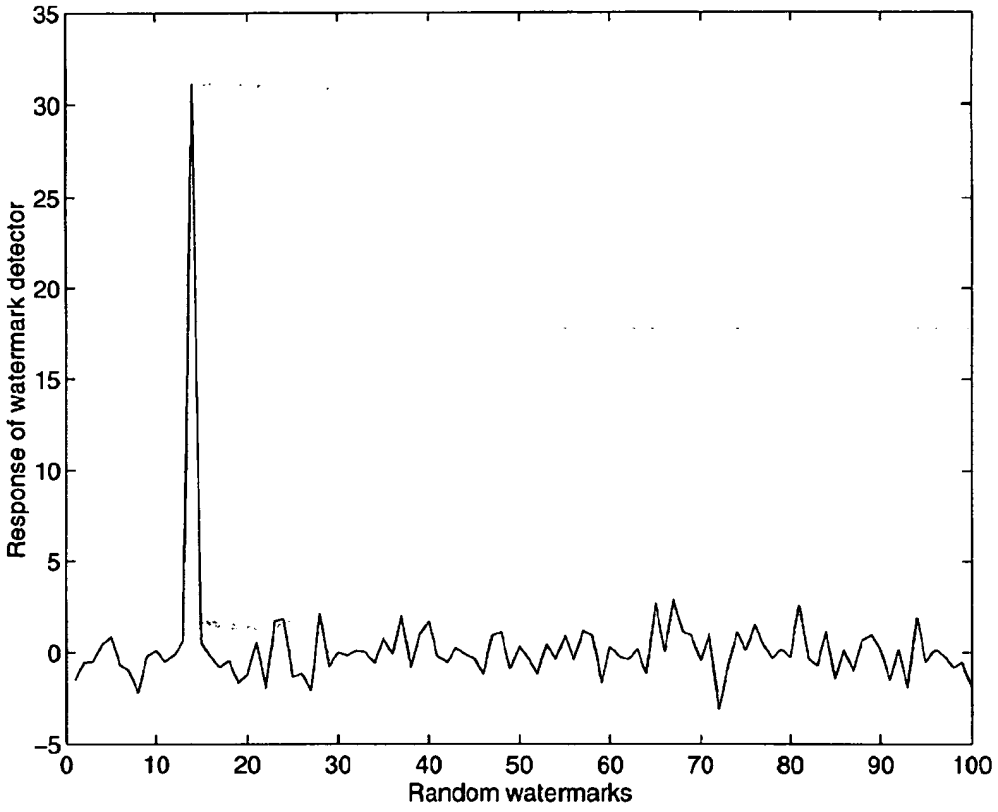

Response of watermark detector
Random watermarks
35
30
25
20
15
10
5
0
−5
0
10
20
30
40
50
60
70
80
90
100

FIGURE 4

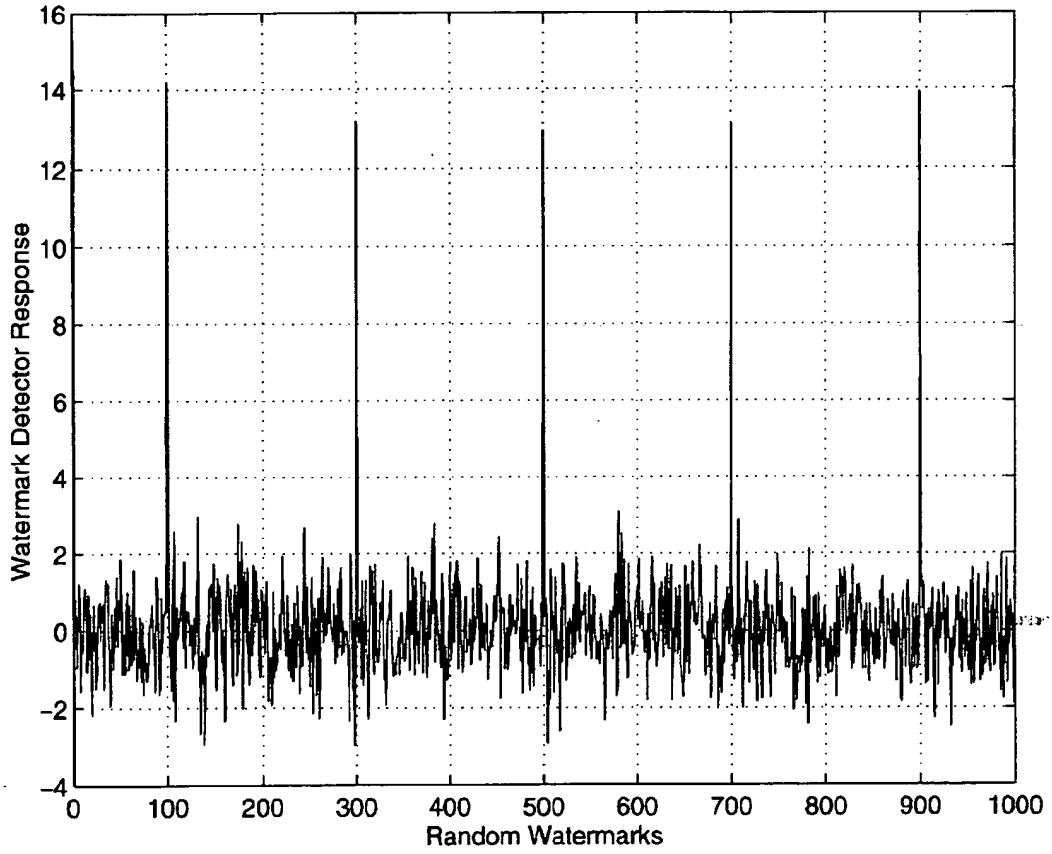

Watermark Detector Response
16
14
12
10
8
6
4
2
0
-2
-4
0
100
200
300
400
500
600
700
800
900
1000
Random Watermarks

FIGURE 5

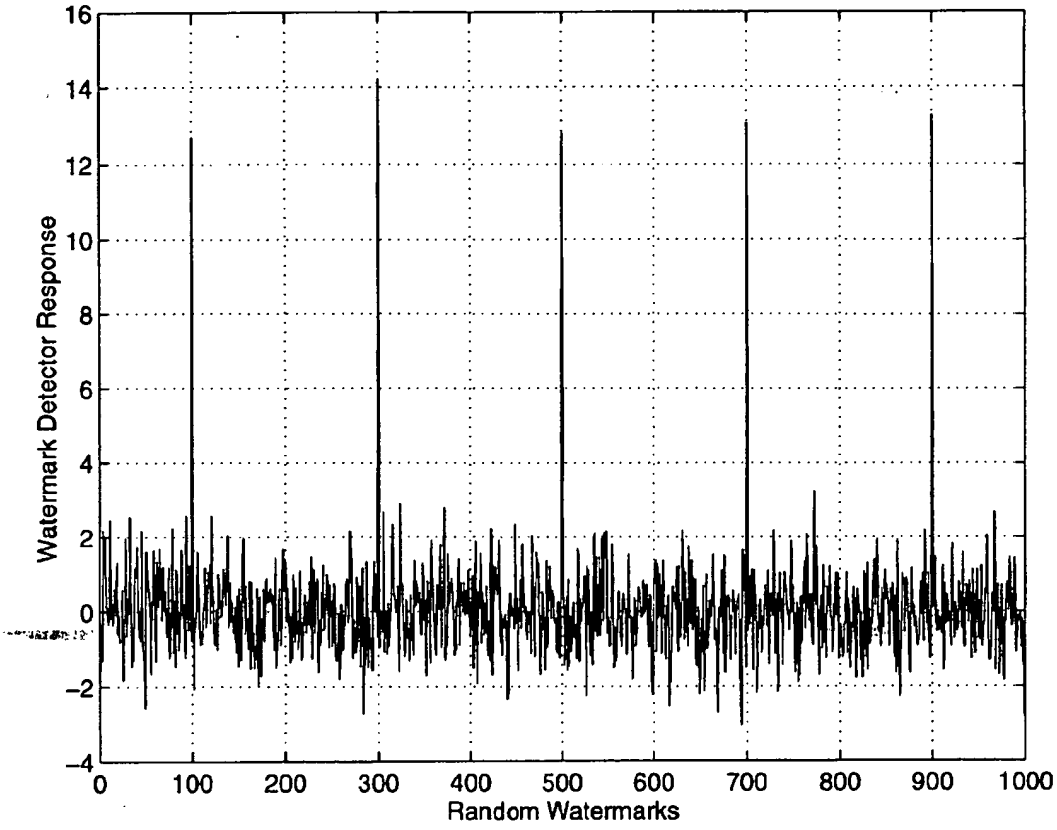

Watermark Detector Response
16
14
12
10
8
6
4
2
0
−2
−4
0
100
200
300
400
500
600
700
800
900
1000
Random Watermarks

FIGURE 6

SECURE SPREAD SPECTRUM WATERMARKING FOR MULTIMEDIA DATA

This application is a continuation of application Ser. No. 08/534,894, filed Sep. 28, 1995, now abandoned.

FIELD OF THE INVENTION

The present invention concerns a method of digital watermarking for use in audio, image, video and multimedia data for the purpose of authenticating copyright ownership, identifying copyright infringers or transmitting a hidden message. Specifically, a watermark is inserted into the perceptually most significant components of a decomposition of the data in a manner so as to be virtually imperceptible. More specifically, a narrow band signal representing the watermark is placed in a wideband channel that is the data.

BACKGROUND OF THE INVENTION

The proliferation of digitized media such as audio, image and video is creating a need for a security system which facilitates the identification of the source of the material. The need manifests itself in terms of copyright enforcement and identification of the source of the material.

Using conventional cryptographic systems permits only valid keyholder access to encrypted data, but once the data is encrypted, it is not possible to maintain records of its subsequent representation or transmission. Conventional cryptography therefore provides minimal protection against data piracy of the type a publisher or owner of data or material is confronted with by unauthorized reproduction or distribution of such data or material.

A digital watermark is intended to complement cryptographic processes. The watermark is a visible or preferably an invisible identification code that is permanently embedded in the data. That is, the watermark remains with the data after any decryption process. As used herein the terms data and material will be understood to refer to audio (speech and music), images (photographs and graphics), video (movies or sequences of images) and multimedia data (combinations of the above categories of materials) or processed or compressed versions thereof. These terms are not intended to refer to ASCII representations of text, but do refer to text represented as an image. A simple example of a watermark is a visible "seal" placed over an image to identify the copyright owner. However, the watermark might also contain additional information, including the identity of the purchaser of the particular copy of the image. An effective watermark should possess the following properties:

1. The watermark should be perceptually invisible or its presence should not interfere with the material being protected.

2. The watermark must be difficult (preferably virtually impossible) to remove from the material without rendering the material useless for its intended purpose. However, if only partial knowledge is known, e.g. the exact location of the watermark within an image is unknown, then attempts to remove or destroy the watermark, for instance by adding noise, should result in severe degradation in data fidelity, rendering the data useless, before the watermark is removed or lost.

3. The watermark should be robust against collusion by multiple individuals who each possess a watermarked copy of the data. That is, the watermark should be robust to the combining of copies of the same data set to destroy the watermarks. Also, it must not be possible for colluders to combine each of their images to generate a different valid watermark.

4. The watermark should still be retrievable if common signal processing operations are applied to the data. These operations include, but are not limited to digital-to-analog and analog-to-digital conversion, resampling, requantization (including dithering and recompression) and common signal enhancements to image contrast and color, or audio bass and treble for example. The watermarks in image and video data should be immune from geometric image operations such as rotation, translation, cropping and scaling.

5. The same digital watermark method or algorithm should be applicable to each of the different media under consideration. This is particularly useful in watermarking of multimedia material. Moreover, this feature is conducive to the implementation of video and image/video watermarking using common hardware.

6. Retrieval of the watermark should unambiguously identify the owner. Moreover, the accuracy of the owner identification should degrade gracefully during attack. Several previous digital watermarking methods have been proposed. L. F. Turner in patent number W089/08915 entitled "Digital Data Security System" proposed a method for inserting an identification string into a digital audio signal by substituting the "insignificant" bits of randomly selected audio samples with the bits of an identification code. Bits are deemed "insignificant" if their alteration is inaudible. Such a system is also appropriate for two dimensional data such as images, as discussed in an article by R. G. Van Schyndel et al entitled "A digital watermark" in Intl. Conf. on Image Processing, vol 2, Pages 86–90, 1994. The Turner method may easily be circumvented. For example, if it is known that the algorithm only affects the least significant two bits of a word, then it is possible to randomly flip all such bits, thereby destroying any existing identification code.

An article entitled "Assuring Ownership Rights for Digital Images" by G. Caronni, in Proc. Reliable IT Systems, VIS '95, 1995 suggests adding tags—small geometric patterns-to-digitized images at brightness levels that are imperceptible. While the idea of hiding a spatial watermark in an image is fundamentally sound, this scheme is susceptible to attack by filtering and redigitization. The fainter such watermarks are, the more susceptible they are to such attacks and geometric shapes provide only a limited alphabet with which to encode information. Moreover, the scheme is not applicable to audio data and may not be robust to common geometric distortions, especially cropping. J. Brassil et al in an article entitled "Electronic Marking and Identification Techniques to Discourage Document Copying" in Proc. of Infocom 94, pp 1278–1287, 1994 propose three methods appropriate for document images in which text is common. Digital watermarks are coded by: (1)vertically shifting text lines, (2) horizontally shifting words, or (3) altering text features such as the vertical endlines of individual characters. Unfortunately, all three proposals are easily defeated, as discussed by the authors. Moreover, these techniques are restricted exclusively to images containing text.

An article by K. Tanaka et al entitled "Embedding Secret Information into a Dithered Multi-level Image" in IEEE Military Comm. Conf., pp216–220, 1990 and K. Mitsui et al in an article entitled "Video-Steganography" in IMA Intellectual Property Proc., vl, pp187–206, 1994, describe several watermarking schemes that rely on embedding watermarks that resemble quantization noise. Their ideas hinge on the notion that quantization noise is typically imperceptible to viewers. Their first scheme injects a watermark into an image by using a predetermined data stream to guide level selection in a predictive quantizer. The data stream is chosen so that the resulting watermark looks like quantization noise.

A variation of this scheme is also presented, where a watermark in the form of a dithering matrix is used to dither an image in a certain way. There are several drawbacks to these schemes. The most important is that they are susceptible to signal processing, especially requantization, and geometric attacks such as cropping. Furthermore, they degrade an image in the same way that predictive coding and dithering can.

In Tanaka et al, the authors also propose a scheme for watermarking facsimile data. This scheme shortens or lengthens certain runs of data in the run length code used to generate the coded fax image. This proposal is susceptible to digital-to-analog and analog-to digital conversions. In particular, randomizing the least significant bit (LSB) of each pixel's intensity will completely alter the resulting run length encoding. Tanaka et al also propose a watermarking method for "color-scaled picture and video sequences". This method applies the same signal transform as JPEG (DCT of 8×8 sub-blocks of an image) and embeds a watermark in the coefficient quantization module. While being compatible with existing transform coders, this scheme is quite susceptible to requantization and filtering and is equivalent to coding the watermark in the least significant bits of the transform coefficients.

In a recent paper, by Macq and Quisquater entitled "Cryptology for Digital TV Broadcasting" in Proc. of the IEEE, 83(6), pp944–957, 1995 there is briefly discussed the issue of watermarking digital images as part of a general survey on cryptography and digital television. The authors provide a description of a procedure to insert a watermark into the least significant bits of pixels located in the vicinity of image contours. Since it relies on modifications of the least significant bits, the watermark is easily destroyed. Further, the method is only applicable to images in that it seeks to insert the watermark into image regions that lie on the edge of contours.

W. Bender et al in article entitled "Techniques for Data Hiding" in Proc. of SPIE, v2420, page 40, July 1995, describe two watermarking schemes. The first is a statistical method called "Patchwork". Patchwork randomly chooses n pairs of image points (a_i, b_i) and increases the brightness at a_i by one unit while correspondingly decreasing the brightness of b_i. The expected value of the sum of the differences of the n pairs of points is claimed to be 2n, provided certain statistical properties of the image are true. In particular, it is assumed that all brightness levels are equally likely, that is, intensities are uniformly distributed. However, in practice, this is very uncommon. Moreover, the scheme may not be robust to randomly jittering the intensity levels by a single unit, and be extremely sensitive to geometric affine transformations.

The second method is called "texture block coding", where a region of random texture pattern found in the image is copied to an area of the image with similar texture. Autocorrelation is then used to recover each texture region. The most significant problem with this technique is that it is only appropriate for images that possess large areas of random texture. The technique could not be used on images of text, for example. Nor is there a direct analog for audio.

In addition to direct work on watermarking images, there are several works of interest in related areas. E. H. Adelson in U.S. Pat. No. 4,939,515 entitled "Digital Signal Encoding and Decoding Apparatus" describes a technique for embedding digital information in an analog signal for the purpose of inserting digital data into an analog TV signal. The analog signal is quantized into one of two disjoint ranges ({0,2,4 . . }, {1,3,5}, for example) which are selected based on the binary digit to be transmitted. Thus Adelson's method is equivalent to watermark schemes that encode information into the least significant bits of the data or its transform coefficients. Adelson recognizes that the method is susceptible to noise and therefore proposes an alternative scheme wherein a 2×1 Hadamard transform of the digitized analog signal is taken. The differential coefficient of the Hadamard transform is offset by 0 or 1 unit prior to computing the inverse transform. This corresponds to encoding the watermark into the least significant bit of the differential coefficient of the Hadamard transform. It is not clear that this approach would demonstrate enhanced resilience to noise. Furthermore, like all such least significant bit schemes, an attacker can eliminate the watermark by randomization.

U.S. Pat. No. 5,010,405 describes a method of interleaving a standard NTSC signal within an enhanced definition television (EDTV) signal. This is accomplished by analyzing the frequency spectrum of the EDTV signal (larger than that of the NTSC signal) and decomposing it into three sub-bands (L,M,H for low, medium and high frequency respectively). In contrast, the NTSC signal is decomposed into two subbands, L and M. The coefficients, M_k, within the M band are quantized into M levels and the high frequency coefficients, H_k, of the EDTV signal are scaled such that the addition of the H_k signal plus any noise present in the system is less than the minimum separation between quantization levels. Once more, the method relies on modifying least significant bits. Presumably, the mid-range rather than low frequencies were chosen because they are less perceptually significant. In contrast, the method proposed in the present invention modifies the most perceptually significant components of the signal.

Finally, it should be noted that many, if not all, of the prior art protocols are not collusion resistant.

Recently, Digimarc Corporation of Portland, Oreg., has described work referred to as signature technology for use in identifying digital intellectual property. Their method adds or subtracts small random quantities from each pixels. Addition or subtraction is based on comparing a binary mask of N bits with the least significant bit (LSB) of each pixel. If the LSB is equal to the corresponding mask bit, then the random quantity is added, otherwise it is subtracted. The watermark is extracted by first computing the difference between the original and watermarked images and then by examining the sign of the difference, pixel by pixel, to determine if it corresponds to the original sequence of additions/subtractions. The Digimarc technique is not based on direct modifications of the image spectrum and does not make use of perceptual relevance. While the technique appears to be robust, it may be susceptible to constant brightness offsets and to attacks based on exploiting the high degree of local correlation present in an image. For example, randomly switching the position of similar pixels within a local neighborhood may significantly degrade the watermark without damaging the image.

In a paper by Koch, Rindfrey and Zhao entitled "Copyright Protection for Multimedia Data", two general methods for watermarking images are described. The first method partitions an image into 8×8 blocks of pixels and computes the Discrete Cosine Transform (DCT) of each of these blocks. A pseudorandom subset of the blocks is chosen and in each such block a triple of frequencies selected from one of 18 predetermined triples is modified so that their relative strengths encode a 1 or 0 value. The 18 possible triples are composed by selection of three out of eight predetermined frequencies within the 8×8 DCT block. The choice of the

eight frequencies to be altered within the DCT block appears to be based on the belief that middle frequencies have a moderate variance level, i.e., they have similar magnitude. This property is needed in order to allow the relative strength of the frequency triples to be altered without requiring a modification that would be perceptually noticeable. Unlike in the present invention, the set of frequencies is not chosen based on any perceptual significance or relative energy considerations. In addition, because the variance between the eight frequency coefficients is small, one would expect that the technique may be sensitive to noise or distortions. This is supported by the experimental results reported in the Koch et al paper, supra, where it is reported that the "embedded labels are robust against JPEG compression for a quality factor as low as about 50%". In contrast, the method described in accordance with the teachings of the present invention has been demonstrated with compression quality factors as low as 5 percent.

An earlier proposal by Koch and Zhao in a paper entitled "Toward Robust and Hidden Image Copyright Labeling" proposed not triples of frequencies but pairs of frequencies and was again designed specifically for robustness to JPEG compression. Nevertheless, the report states that "a lower quality factor will increase the likelihood that the changes necessary to superimpose the embedded code on the signal will be noticeably visible".

In a second method, proposed by Koch and Zhao, designed for black and white images, no frequency transform is employed. Instead, the selected blocks are modified so that the relative frequency of white and black pixels encodes the final value. Both watermarking procedures are particularly vulnerable to multiple document attacks. To protect against this, Zhao and Koch proposed a distributed 8×8 block of pixels created by randomly sampling 64 pixels from the image. However, the resulting DCT has no relationship to that of the true image. Consequently, one would expect such distributed blocks to be both sensitive to noise and likely to cause noticeable artifacts in the image.

In summary, prior art digital watermarking techniques are not robust and the watermark is easy to remove. In addition, many prior techniques would not survive common signal and geometric distortions

SUMMARY OF THE INVENTION

The present invention overcomes the limitations of the prior art methods by providing a watermarking system that embeds an unique identifier into the perceptually significant components of a decomposition of an image, an audio signal or a video sequence.

Preferably, the decomposition is a spectral frequency decomposition. The watermark is embedded in the data's perceptually significant frequency components. This is because an effective watermark cannot be located in perceptually insignificant regions of image data or in its frequency spectrum, since many common signal or geometric processes affect these components. For example, a watermark located in the high frequency spectral components of an image is easily removed, with minor degradation to the image, by a process that performs low pass filtering. The issue then becomes one of how to insert the watermark into the most significant regions of the data frequency spectrum without the alteration being noticeable to an observer, i.e., a human or a machine feature recognition system. Any spectral component may be altered, provided the alteration is small. However, very small alterations are susceptible to any noise present or intentional distortion.

In order to overcome this problem, the frequency domain of the image data or sound data may be considered as a communication channel, and correspondingly the watermark may be considered as a signal transmitted through the channel. Attacks and intentional signal distortions are thus treated as noise from which the transmitted signal must be immune. Attacks are intentional efforts to remove, delete or otherwise overcome the beneficial aspects of the data watermarking. While the present invention is intended to embed watermarks in data, the same methodology can be applied to sending any type of message through media data.

Instead of encoding the watermark into the least significant components of the data, the present invention considers applying concepts of spread spectrum communication. In spread spectrum communications, a narrowband signal is transmitted over a much larger bandwidth such that the signal energy present in any single frequency is imperceptible. In a similar manner, the watermark is spread over many frequency bins so that the energy in any single bin is small and imperceptible. Since the watermark verification process includes a priori knowledge of the locations and content of the watermarks, it is possible to concentrate these many weak signals into a single signal with a high signal to-noise ratio. Destruction of such a watermark would require noise of high amplitude to be added to every frequency bin.

In accordance with the teachings of the present invention, a watermark is inserted into the perceptually most significant regions of the data decomposition. The watermark itself is designed to appear to be additive random noise and is spread throughout the image. By placing the watermark into the perceptually significant components, it is much more difficult for an attacker to add more noise to the components without adversely affecting the image or other data. It is the fact that the watermark looks like noise and is spread throughout the image or data which makes the present scheme appear to be similar to spread spectrum methods used in communications system.

Spreading the watermark throughout the spectrum of an image ensures a large measure of security against unintentional or intentional attack. First, the location of the watermark is not obvious. Second, frequency regions are selected in a fashion that ensures severe degradation of the original data following any attack on the watermark.

A watermark that is well placed in the frequency domain of an image or a sound track will be practically impossible to see or hear. This will always be the case if the energy in the watermark is sufficiently small in any single frequency coefficient. Moreover, it is possible to increase the energy present in particular frequencies by exploiting knowledge of masking phenomena in the human auditory and visual systems. Perceptual masking refers to any situation where information in certain regions of an image or a sound is occluded by perceptually more prominent information in another part of the image or sound. In digital waveform coding, this frequency domain (and in some cases, time/pixel domain) masking is exploited extensively to achieve low bit rate encoding of data. It is clear that both auditory and visual systems attach more resolution to the high energy, low frequency, spectral regions of an auditory or visual scene. Further, spectrum analysis of images and sounds reveals that most of the information in such data is often located in the low frequency regions.

In addition, particularly for processed or compressed data, perceptually significant need not refer to human perceptual significance, but may refer instead to machine perceptual significance, for instance, machine feature recognition.

To meet these requirements, a watermark is proposed whose structure comprises a large quantity, for instance 1000, of randomly generated numbers with a normal distribution having zero mean and unity variance. A binary watermark is not chosen because it is much less robust to attacks based on collusion of several independently watermarked copies of an image. However, generally, the watermark might have arbitrary structure, both deterministic and/or random, and including uniform distributions. The length of the proposed watermark is variable and can be adjusted to suit the characteristics of the data. For example, longer watermarks might be used for images that are especially sensitive to large modifications of its spectral coefficients, thus requiring weaker scaling factors for individual components.

The watermark is then placed in components of the image spectrum. These components may be chosen based on an analysis of those components which are most vulnerable to attack and/or which are most perceptually significant. This ensures that the watermark remains with the image even after common signal and geometric distortions. Modification of these spectral components results in severe image degradation long before the watermark itself is destroyed. Of course, to insert the watermark, it is necessary to alter these very same coefficients. However, each modification can be extremely small and, in a manner similar to spread spectrum communication, a strong narrowband watermark may be distributed over a much broader image (channel) spectrum. Conceptually, detection of the watermark then proceeds by adding all of these very small signals, whose locations are only known to the copyright owner, and concentrating the watermark into a signal with high signal-to-noise ratio. Because the location of the watermark is only known to the copyright holder, an attacker would have to add very much more noise energy to each spectral coefficient in order to be confident of removing the watermark. However, this process would destroy the image.

Preferably, a predetermined number of the largest coefficients of the DCT (discrete cosine transform) (excluding the DC term) are used. However, the choice of the DCT is not critical to the algorithm and other spectral transforms, including wavelet type decompositions are also possible. In fact, use of the FFT rather than DCT is preferable from a computational perspective.

The invention will be more clearly understood when the following description is read in conjunction with the accompanying drawing.

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 is a schematic representation of typical common processing operations to which data could be subjected;

FIG. 2 is a schematic representation of a preferred system for immersing a watermark into an image;

FIGS. 3*a* and 3*b* are flow charts of the encoding and decoding of watermarks;

FIG. 4 is a graph of the responses of the watermark detector to random watermarks;

FIG. 5 is a graph of the response of the watermark detector to random watermarks for an image which is successively watermarked five times;

FIG. 6 is a graph of the response of the watermark detector to random watermarks where five images, each having a different watermark, and averaged together; and

FIG. 7 is a schematic diagram of an optical embodiment of the present invention

DETAILED DESCRIPTION

In order to better understand the advantages of the invention, the preferred embodiment of a frequency spectrum based watermarking system will be described. It is instructive to examine the processing stages that image (or sound) data may undergo in the copying process and to consider the effect that such processing stages can have on the data. Referring to FIG. 1, a watermarked image or sound data **10** is transmitted **12** to undergo typical distortion or intentional tampering **14**. Such distortions or tampering includes lossy compression **16**, geometric distortion **18**, signal processing **20** and D/A and A/D conversion **22**. After undergoing distortion or tampering, corrupted watermarked image or sound data **24** is transmitted **26**. The process of "transmission" refers to the application of any source or channel code and/or of encryption techniques to the data. While most transmission steps are information lossless, many compression schemes (e.g., JPEG, MPEG, etc.) may potentially degrade the quality of the data through irretrievable loss of data. In general, a watermarking method should be resilient to any distortions introduced by transmission or compression algorithms.

Lossy compression **16** is an operation that usually eliminates perceptually irrelevant components of image or sound data. In order to preserve a watermark when undergoing lossy compression, the watermark is located in a perceptually significant region of the data. Most processing of this type occurs in the frequency domain. Data loss usually occurs in the high frequency components. Thus, the watermark must be placed in the significant frequency component of the image (or sound) data spectrum to minimize the adverse affects of lossy compression.

After receipt, an image may encounter many common transformations that are broadly categorized as geometric distortions or signal distortions. Geometric distortions 18 are specific to image and video data, and include such operations as rotation, translation, scaling and cropping. By manually determining a minimum of four or nine corresponding points between the original and the distorted watermark, it is possible to remove any two or three dimensional affine transformation. However, an affine scaling (shrinking) of the image results in a loss of data in the high frequency spectral regions of the image. Cropping, or the cutting out and removal of portions of an image, also results in irretrievable loss of data. Cropping may be a serious threat to any spatially based watermark but is less likely to affect a frequency-based scheme.

Common signal distortions include digital-to-analog and analog-to-digital conversion 22, resampling, requantization, including dithering and recompression, and common signal enhancements to image contrast and/or color, and audio frequency equalization. Many of these distortions are non-linear, and it is difficult to analyze their effect in either a spatial or frequency based method. However, the fact that the original image is known allows many signal transformations to be undone, at least approximately. For example, histogram equalization, a common non-linear contrast enhancement method, may be substantially removed by histogram specification or dynamic histogram warping techniques.

Finally, the copied image may not remain in digital form. Instead, it is likely to be printed or an analog recording made (analog audio or video tape). These reproductions introduce additional degradation into the image data that a watermarking scheme must be robust to.

Tampering (or attack) refers to any intentional attempt to remove the watermark, or corrupt it beyond recognition. The

watermark must not only be resistant to the inadvertent application of distortions. It must also be immune to intentional manipulation by malicious parties. These manipulations can include combinations of distortions, and can also include collusion and forgery attacks.

FIG. 2 shows a preferred system for inserting a watermark into an image in the frequency domain. Image data $X(i,j)$ assumed to be in digital form, or alternatively data in other formats such as photographs, paintings or the like, that have been previously digitized by well-known methods, is subject to a frequency transformation **30**, such as the Fourier transform. A watermark signal W (k) is inserted into the frequency spectrum components of the transformed image data **32** applying the techniques described below. The frequency spectrum image data including the watermark signal is subjected to an inverse frequency transform **34**, resulting in watermarked image data $\hat{X}(i,j)$, which may remain in digital form or be printed as an analog representation by well-known methods.

After applying a frequency transformation to the image data **30**, a perceptual mask is computed that highlights prominent regions in the frequency spectrum capable of supporting the watermark without overly affecting perceptual fidelity. This may be performed by using knowledge of the perceptual significance of each frequency in the spectrum, as discussed earlier, or simply by ranking the frequencies based on their energy. The latter method was used in experiments described below.

In general, it is desired to place the watermark in regions of the spectrum that are least affected by common signal distortions and are most significant to image quality as perceived by a viewer, such that significant modification would destroy the image fidelity. In practice, these regions could be experimentally identified by applying common signal distortions to images and examining which frequencies are most affected, and by psychophysical studies to identify how much each component may be modified before significant changes in the image are perceivable.

The watermark signal is then inserted into these prominent regions in a way that makes any tampering create visible (or audible) defects in the data. The requirements of the watermark mentioned above and the distortions common to copying provide constraints on the design of an electronic watermark.

In order to better understand the watermarking method, reference is made to FIGS. 3(*a*) and 3(*b*) where from each document D a sequence of values $X=x_1, \ldots, x_n$ is extracted **40** with which a watermark $W=w_1, \ldots, w_n$ is combined **42** to create an adjusted sequence of values $X'=x'_1, \ldots, x'_n$ which is then inserted back **44** into the document in place of values X in order to obtain a watermark document D'. An attack of the document D', or other distortion, will produce a document D*. Having the original document D and the document D*, a possibly corrupted watermark W* is extracted **46** and compared to watermark W **48** for statistical analysis **50**. The values W* are extracted by first extracting a set of values $X^*=x_1^*, \ldots, x_n^*$ from D* (using information about D) and then generating W* from the values X* and the values X.

When combining the values X with the watermark values W in step **42**, scaling parameter α is specified. The scaling parameter α determines the extent to which values W alter values X. Three preferred formulas for computing X' are:

$$x_i'=x_i+\alpha w_i \quad (1)$$

$$x_i'=x_i(1+\alpha w_i) \quad (2)$$

$$x_i'=x_i(e^{\alpha w_i}) \quad (3)$$

Equation 1 is invertible. Equations 2 and 3 are invertible when $x_i \neq 0$. Therefore, given X* it is possible to compute the inverse function necessary to derive W* from X and X*.

Equation 1 is not the preferred formula when the values x_i vary over a wide range. For example, if $x_i=10^6$ then adding 100 may be insufficient to establish a watermark, but if $x_i=10$, then adding 100 will unacceptably distort the value. Insertion methods using equations 2 and 3 are more robust when encountering such a wide range of values x_i. It will also be observed that equation 2 and 3 yield similar results when αw_i is small. Moreover, when x_i is positive, equation 3 is equivalent to $\ln(x_i')=\ln(x_i)+\alpha x_i$ and may be considered as an application of equation 1 when natural logarithms of the original values are used. For example, if $|w_i| \leqq 1$ and $\alpha=0.01$, then using Equation (2) guarantees that the spectral coefficient will change by no more than 1%.

For certain applications, a single scaling parameter α may not be best for combining all values of x_i. Therefore, multiple scaling parameters $\alpha_i, \ldots, \alpha_n$ can be used with revised equations 1 to 3 such as $x_i'=x_i(1+\alpha_i w_i)$. The values of α_i serve as a relative measure of how much x_i must be altered to change the perceptual quality of the document. A large value for α_i means that it is possible to alter x_i by a large amount without perceptually degrading the document.

A method for selecting the multiple scaling values is based upon certain general assumptions. For example, equation 2 is a special case of the generalized equation 1, $(x_i'=x_i+\alpha_i x_i)$, for $\alpha_i=\alpha x_i$. That is, equation 2 makes the reasonable assumption that a large value of x_i is less sensitive to additive alteration that a small value of x_i.

Generally, the sensitivity of the image to different values of α_i is unknown. A method of empirically estimating the sensitivities is to determine the distortion caused by a number of attacks on the original image. For example, it is possible to compute a degraded image D* from D, extract the corresponding values $x_1^*, \ldots, x_n^*$ and select α_i to be proportional to the deviation $|x_i^*-x_i|$. For greater robustness, it is possible to try other forms of distortion and make α_i proportional to the average value of $|x_i^*-x_i|$. Instead of using the average distortion, it is possible to use the median or maximum deviation.

Alternatively, it is possible to combine the empirical approach with general global assumptions regarding the sensitivity of the values. For example, it might be required that $\alpha_i \geqq \alpha_j$ whenever $x_i \geqq x_j$. This can be combined with the empirical approach by setting α_i according to

$$\alpha_i \sim \max_{\{j \mid v_j \leq v_i\}} |v_j^* - v_j|$$

A more sophisticated approach is to weaken the monotonicity constraint to be robust against occasional outliers.

The length of the watermark, n, determines the degree to which the watermark is spread among the relevant components of the image data. As the size of the watermark increases, so does the number of altered spectral components, and the extent to which each component need be altered decreases for the same resilience to noise. Consider watermarks of the form $x_i'=x_i+\alpha w_i$ and a white noise attack by $x_i''=x_i'+r_i$, where r_i are chosen according to independent normal distributions with standard deviation σ. It is possible to recover the watermark when α is proportional to $\sigma/\sqrt{n}$. That is, quadrupling the number of components can halve the magnitude of the watermark placed into each component. The sum of the squares of the deviations remains essentially unchanged.

In general, a watermark comprises an arbitrary sequence of real numbers $W=w_1, \ldots, w_n$. In practice, each value w_i may be chosen independently from a normal distribution N(0,1), where $N(\mu, \sigma^2)$ with mean μ and variance σ^2 or of a uniform distribution from {1,−1} or {0,1}.

It is highly unlikely that the extracted mark W* will be identical to the original watermark W. Even the act of requantizing the watermarked document for transmission will cause W* to deviate from W. A preferred measure of the similarity of W and W* is

$$sim(W, W^*) = \frac{W^* \cdot W}{\sqrt{W^* \cdot W^*}} \quad (4)$$

Large values of sim (W,W*) are significant in view of the following analysis. Assume that the authors of document D* had no access to W (either through the seller or through a watermarked document). Then for whatever value of W* is obtained, the conditional distribution on w_i will be independently distributed according to N(0,1). In this case,

$$N\left(0, \sum_{i=1}^{n} x_i^2\right) = N(0, W^* \cdot W^*).$$

Thus, sim(W,W*) is distributed according to N(0,1). Then, one may apply the standard significance tests for the normal distribution. For example, if D* is chosen independently from W, then it is very unlikely that sim(W,W*)>5. Note that somewhat higher values of sim (W,W*) may be needed when a large number of watermarks are on file. The above analysis required only the independence of W from W*, and did not rely on any specific properties of W* itself. This fact provides further flexibility when preprocessing W*.

The extracted watermark W* may be extracted in several ways to potentially enhance the ability to extract a watermark. For example, experiments on images encountered instances where the average value of W*, denoted $E_i(W^*)$, differed substantially from 0, due to the effects of a dithering procedure. While this artifact could be easily eliminated as part of the extraction process, it provides a motivation for postprocessing extracted watermarks. As a result, it was discovered that the simple transformation $w_i^* \leftarrow w_i^* - E_i(W^*)$ yielded superior values of sim (W,W*). The improved performance resulted from the decreased value of W*. W*; the value of W*. W was only slightly affected.

In experiments it was frequently observed that w_i^* could be greatly distorted for some values of i. One postprocessing option is to simply ignore such values, setting them to 0. That is,

$$w_i^* \leftarrow \begin{cases} w_i^* & \text{if } |w_i^*| > \text{tolerance} \\ 0 & \text{otherwise} \end{cases}$$

The goal of such a transformation is to lower W*·W*. A less abrupt version of this approach is to normalize the W* values to be either −1,0 or 1, by

$$w_i^* \leftarrow \text{sign}(w_i^* - E_i(W^*)).$$

This transformation can have a dramatic effect on the statistical significance of the result. Other robust statistical techniques could also be used to suppress outlier effects.

In principle, any frequency domain transform can be used. In the scheme described below, a Fourier domain method is used, but the use of wavelet based schemes are also useable as a variation. In terms of selecting frequency regions of the transform, it is possible to use models for the perceptual system under consideration.

Frequency analysis may be performed by a wavelet or sub-band transform where the signal is divided into sub-bands by means of a wavelet or multi-resolution transform. The sub-bands need not be uniformly spaced. Each sub-band may be thought of as representing a frequency region in the domain corresponding to a sub-region of the frequency range of the signal. The watermark is then inserted into the sub-regions.

For audio data, a sliding "window" moves along the signal data and the frequency transform (DCT, FFT, etc.) is taken of the sample in the window. This process enables the capture of meaningful information of a signal that is time varying in nature.

Each coefficient in the frequency domain is assumed to have a perceptual capacity. That is, it can support the insertion of additional information without any (or with minimal) impact to the perceptual fidelity of the data.

In order to place a length L watermark into an N×N image, the N×N FFT (or DCT) of the image is computed and the watermark is placed into the L highest magnitude coefficients of the transform matrix, excluding the DC component. More generally, L randomly chosen coefficients could be chosen from the M, M≧L most perceptually significant coefficients of the transform. For most images, these coefficients will be the ones corresponding to the low frequencies. The purpose of placing the watermark in these locations is because significant tampering with these frequencies will destroy the image fidelity or perceived quality well before the watermark is destroyed.

The FFT provides perceptually similar results to the DCT. This is different than the case of transform coding, where the DCT is preferred to the FFT due to its spectral properties. The DCT tends to have less high frequency information than that the FFT, and places most of the image information in the low frequency regions, making it preferable in situations where data need to be eliminated. In the case of watermarking, image data is preserved, and nothing is eliminated. Thus the FFT is as good as the DCT, and is preferred since it is easier to compute.

In an experiment, a visually imperceptible watermark was intentionally placed in an image. Subsequently, 100 randomly generated watermarks, only one of which corresponded to the correct watermark, were applied to the watermark detector described above. The result, as shown in FIG. 4, was a very strong positive response corresponding to the correct watermark, suggesting that the method results in a very low number of false positive responses and a very low false negative response rate.

In another test, the watermarked image was scaled to half of its original size. In order to recover the watermark, the image was re-scaled to its original size, albeit with loss of detail due to subsampling of the image using low pass spatial filter operations. The response of the watermark detector was well above random chance levels, suggesting that the watermark is robust to geometric distortions. This result was achieved even though 75 percent of the original data was missing from the scaled down image.

In a further experiment, a JPEG encoded version of the image with parameters of 10 percent quality and 0 percent smoothing, resulting in visible distortions, was used. The results of the watermark detector suggest that the method is robust to common encoding distortions. Even using a version of the image with parameters of the 5 percent quality

and 0 percent smoothing, the results were well above that achievable due to random chance.

In experiments using a dithered version of the image, the response of the watermark detector suggested that the method is robust to common encoding distortion. Moreover, more reliable detection is achieved by removing any non-zero mean from the extracted watermark.

In another experiment, the image was clipped, leaving only the central quarter of the image. In order to extract the watermark from the clipped image, the missing portion of the image was replaced with portions from the original unwatermarked image. The watermark detector was able to recover the watermark with a response greater than random. When the non-zero mean was removed, and the elements of the watermark were binarized prior to the comparison with the correct watermark, the detector response was improved. This result is achieved even though 75 percent of the data was removed from the image.

In yet another experiment, the image was printed, photocopied, scanned using a 300 dpi Umax PS-2400x scanner and rescaled to a size of 256×256 pixels. Clearly, the final image suffered from different levels of distortion introduced at each process. High frequency pattern noise was particularly noticeable. When the non-zero mean was removed and only the sign of the elements of the watermark was used, the watermark detector response improved to well above random chance levels.

In still another experiment, the image was subject to five successive watermarking operations. That is, the original image was watermarked, the watermarked image was watermarked, and so forth. The process may be considered another form of attack in which it is clear that significant image degradation occurs if the process is repeated. FIG. **5** shows the response of the watermark detector to 1000 randomly generated watermarks, including the five watermarks present in the image. The five dominant spikes in the graph, indicative of the presence of the five watermarks, show that successive watermarking does not interfere with the process.

The fact that successive watermarking is possible means that the history or pedigree of a document is determinable if successive watermarking is added with each copy.

In a variation of the multiple watermark image, five separately watermarked images were averaged together to simulate simple conclusion attack. FIG. **6** shows the response of the watermark detector to 1000 randomly generated watermarks, including the five watermarks present in the original images. The result is that simple collusion based on averaging is ineffective in defeating the present watermarking system.

The result of the above experiments is that the described system can extract a reliable copy of the watermark from images that have been significantly degraded through several common geometric and signal processing procedures. These procedures include zooming (low pass filtering), cropping, lossy JPEG encoding, dithering, printing, photocopying and subsequent rescanning.

While these experiments were, in fact, conducted using an image, similar results are attainable with text images, audio data and video data, although attention must be paid to the time varying nature of these data.

The above implementation of the watermarking system is an electronic system. Since the basic principle of the invention is the inclusion of a watermark into spectral frequency components of the data, watermarking can be accomplished by other means using, for example, an optical system as shown in FIG. **7**.

In FIG. **7**, data to be watermarked such as an image **52** is passed through a spatial transform lens **54**, such as a Fourier transform lens, the output of which lens is the spatial transform of the image. Concurrently, a watermark image **56** is passed through a second spatial transform lens **58**, the output of which lens is the spatial transfer of the watermark image **56**. The spatial transform from lens **54** and the spatial transform from lens **58** are combined at an optical combiner **60**. The output of the optical combiner **60** is passed through an inverse spatial transform lens **62** from which the watermark image **64** is present. The result is a unique, virtually imperceptible, watermarked image. Similar results are achievable by transmitting video or multimedia signals through the lenses in the manner described above.

While there have been described and illustrated spread spectrum watermarking of data and variations and modifications thereof, it will be apparent to those skilled in the art that further variations and modifications are possible without deviating from the broad principles and spirit of the present invention which shall be limited solely by the scope of the claims appended hereto.

What is claimed is:

1. A method of inserting a watermark into data comprising the steps of:

obtaining a spectral decomposition of data to be watermarked which data is a representation of humanly perceivable material;

inserting a watermark into the perceptually significant components of the decomposition of data: and

applying an inverse transform to the decomposition of data with the watermark for generating watermarked data.

2. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises image data.

3. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises video data.

4. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises audio data.

5. A method of inserting a watermark into data as set forth in claim **1**, where said data comprises multimedia data.

6. A method of inserting a watermark into data as set forth in claim **1**, where said obtaining a spectral decomposition of data is selected from the group consisting of Fourier transformation, discrete cosine transformation, Hadamard transformation, and wavelet, multi-resolution, sub-band method.

7. A method of inserting a watermark into data as set forth in claim **6**, where said inserting a watermark inserts watermark values where addition of additional signal into a perceptually significant component affects the perceived quality of the data.

8. A method of inserting a watermark into data as set forth in claim **1**, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

9. The method of inserting a watermark into data as set forth in claim **8**, wheye watermark values include associated scaling parameters.

10. A method of inserting a watermark into data as set forth in claim **9**, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

11. A method of inserting a watermark into data as set forth in claim 8, where the watermark values are chosen according to a random distribution.

12. A method of inserting a watermark into data comprising the steps of:

extracting values of perceptually significant components of a spectral decomposition of data which data is a representation of human perceivable material;

combining watermark values with the extracted values to create adjusted values; and

inserting the adjusted values into the data in place of the extracted values to produce watermarked data.

13. The method of inserting a watermark into data as set forth in claim 12, where watermark values include associated scaling parameters.

14. A method of inserting a watermark into data as set forth in claim 13, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

15. A method of inserting a watermark into data as set forth in claim 12, where the watermark values are chosen according to a random distribution.

16. A method of inserting a watermark into data as set forth in claim 12, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

17. The method of inserting a watermark into data as set forth in claim 16, where watermark values include associated scaling parameters.

18. A method of inserting a watermark into data as set forth in claim 12, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

19. A method of inserting a watermark into data as set forth in claim 16, where the watermark values are chosen according to a random distribution.

20. A method of inserting a watermark into data as set forth in claim 16, further comprising the step of preprocessing distorted or tampered watermarked data before said comparing data.

21. A method of inserting a watermark into data as set forth in claim 20, where said distorted or tampered watermarked data is clipped data and said preprocessing comprises replacing missing portions of the data with corresponding portions from original unwatermarked data.

22. A method of inserting a watermark into data as set forth in claim 12, where said combining watermark values sequentially combines watermark values for a plurality of watermarks.

23. A system for inserting a watermark into data comprising:

providing image data;

providing watermark data;

first transform lens for transforming image data passing therethrough into transformed image data;

second transform lens for transforming watermark data passing therethrough into transformed watermark data;

optical combiner for combining the transformed image data and the transformed watermark data to form transformed watermarked data; and

inverse transform lens for forming watermarked data by inverse transformation of transformed watermarked data.

24. A system for inserting a watermark into data as set forth in claim 23, where said first transform lens and said second transform lens are Fourier transform lenses and said inverse transform lens is an inverse Fourier transform lens.

25. A method of inserting a watermark into data comprising the steps of:

providing a medium containing data;

obtaining a spectral decomposition of data to be watermarked;

inserting a watermark into the perceptually significant components of the decomposition of data; and

applying an inverse transform to the decomposition of data with the watermark to generate watermarked data.

26. A method of inserting a watermark into data as set forth in claim 25, where said data comprises image data.

27. A method of inserting a watermark into data as set forth in claim 25, where said data comprises video data.

28. A method of inserting a watermark into data as set forth in claim 25, where said data comprises audio data.

29. A method of inserting a watermark into data as set forth in claim 25, where said data comprises multimedia data.

30. A method of inserting a watermark into data as set forth in claim 25, where said obtaining a spectral decomposition of data is selected from the group consisting of Fourier transformation, discrete cosine transformation, Hadamard transformation, and wavelet, multi-resolution, subband method.

31. A method of inserting a watermark into data as set forth in claim 30, where said inserting a watermark inserts watermark values where addition of additional signal into a perceptually significant component affects the perceived quality of the data.

32. A method of inserting a watermark into data as set forth in claim 25, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

33. The method of inserting a watermark into data as set forth in claim 32, where watermark values include associated scaling parameters.

34. A method of inserting a watermark into data as set forth in claim 33, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

35. A method of inserting a watermark into data as set forth in claim 32, where the watermark values are chosen according to a random distribution.

36. A method of inserting a watermark into data comprising the steps of:

providing a medium containing data;

extracting values of perceptually significant components of a spectral decomposition of the data;

combining watermark values with the extracted values to create adjusted values; and

inserting the adjusted values into the data in place of the extracted values to produce watermarked data.

37. The method of inserting a watermark into data as set forth in claim 36, where watermark values include associated scaling parameters.

38. A method of inserting a watermark into data as set forth in claim 37, where scaling parameters are selected such

that adding additional watermark value affects the perceived quality of the data.

39. A method of inserting a watermark into data as set forth in claim 36, where the watermark values are chosen according to a random distribution.

40. A method of inserting a watermark into data as set forth in claim 36, further comprising:

comparing data with watermarked data for obtaining extracted data values;

comparing extracted data values with watermark values and data for obtaining difference values; and

analyzing difference values to determine the watermark in the watermarked data.

41. The method of inserting a watermark into data as set forth in claim 40, where watermark values include associated scaling parameters.

42. A method of inserting a watermark into data as set forth in claim 41, where scaling parameters are selected such that adding additional watermark value affects the perceived quality of the data.

43. A method of inserting a watermark into data as set forth in claim 40, where the watermark values are chosen according to a random distribution.

44. A method of inserting a watermark into data as set forth in claim 40, further comprising the step of preprocessing distorted or tampered watermarked data before said comparing data.

45. A method of inserting a watermark into data as set forth in claim 44, where said distorted or tampered watermarked data is clipped data and said preprocessing comprises replacing missing portions of the data with corresponding portions from original unwatermarked data.

46. A method of inserting a watermark into data as set forth in claim 36, where said combining watermark values sequentially combines watermark values for a plurality of watermarks.

* * * * *

US006226387B1

(12) **United States Patent**
Tewfik et al.

(10) **Patent No.: US 6,226,387 B1**
(45) **Date of Patent: May 1, 2001**

(54) **METHOD AND APPARATUS FOR SCENE-BASED VIDEO WATERMARKING**

(75) Inventors: **Ahmed H. Tewfik**, Edina; **Mitchell D. Swanson**, Minneapolis; **Bin Zhu**, St. Paul, all of MN (US)

(73) Assignee: **Regents of the University of Minnesota**, Minneapolis, MN (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **08/921,931**

(22) Filed: **Aug. 27, 1997**

Related U.S. Application Data

(60) Provisional application No. 60/050,587, filed on Jun. 24, 1997, and provisional application No. 60/024,979, filed on Aug. 30, 1996.

(51) **Int. Cl.**[7] **G06K 9/00**

(52) **U.S. Cl.** **382/100**

(58) **Field of Search** 382/100, 232; 380/210, 287, 54

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,395,024	7/1968	Earle	99/169
4,313,197	1/1982	Maxemchuk	370/111
4,425,661	1/1984	Moses et al.	375/1
4,495,620	1/1985	Steele et al.	370/118
4,969,041	11/1990	O'Grady et al.	358/142
5,010,405	4/1991	Schreiber et al.	358/141
5,060,262	10/1991	Bevins, Jr. et al.	380/19
5,285,498	2/1994	Johnston	381/2
5,315,098	5/1994	Tow	235/494
5,319,735	6/1994	Peuss et al.	395/2.14
5,325,290	6/1994	Cauffman et al.	364/401
5,379,345	1/1995	Greenberg	380/23
5,386,240	1/1995	Hori	348/473
5,404,377	4/1995	Moses	375/200
5,450,490	9/1995	Jensen et al.	380/6
5,461,426	10/1995	Limberg et al.	348/475
5,465,269	11/1995	Schaffner et al.	375/200
5,465,308	11/1995	Hutcheson et al.	382/159
5,473,631	12/1995	Moses	375/202
5,515,296	5/1996	Agarwal	364/514 R
5,530,759	6/1996	Braudaway et al.	380/54

(List continued on next page.)

FOREIGN PATENT DOCUMENTS

0581317A2	2/1994	(EP)	G07D/7/00
0 635 798 A1	1/1995	(EP)	G06F/17/30
0635798A1	1/1995	(EP)	G06F/17/30
0 657 831 A1	6/1995	(EP)	G06F/17/30
0657831A1	6/1995	(EP)	G06F/17/30
07-160731	6/1995	(JP)	G06F/17/30

OTHER PUBLICATIONS

Ohnishi et al., "Embedding a Seal into a Picture under Orthogonal Wavelet Transform," IEEE Proc. 3rd Int. Conf. on Multimedia Computing and Systems, Jun. 17–23, 1996, pp. 514–521.*

Aizawa, K., "Model–Based Image Coding", *Proceedings of the SPIE, Visual Communications and Image Processing '94 vol. 2308*, Chicago, IL, 1035–1049 (Sep. 25–29, 1994).

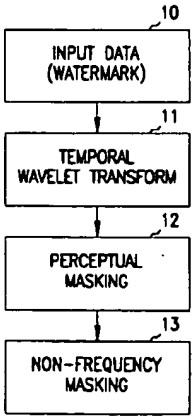
(List continued on next page.)

Primary Examiner—Andrew W. Johns
(74) *Attorney, Agent, or Firm*—Schwegman, Lundberg, Woessner & Kluth P.A.

(57) **ABSTRACT**

A method and apparatus for the scene-based watermarking of video data is disclosed. In one embodiment, each of a number of frames of a scene of video host data undergoes a temporal wavelet transform, from which blocks are extracted. The blocks undergo perceptual masking in the frequency domain, such that a watermark is embedded therein. Once the watermark block is taken out of the frequency domain, a spatial mask of the original block is weighted to the watermark block, and added to the original block to obtain the watermarked block.

14 Claims, 5 Drawing Sheets



U.S. PATENT DOCUMENTS

5,579,471	11/1996	Barber et al.	395/326
5,583,941	12/1996	Yoshida et al.	380/51
5,606,609	2/1997	Houser et al.	380/4
5,613,004	3/1997	Cooperman et al.	380/28
5,646,997	7/1997	Barton	380/23
5,687,236	11/1997	Moskowitz et al.	380/28
5,710,719	1/1998	Houle	364/514 R
5,710,916	1/1998	Barbara et al.	395/609
5,809,139	9/1998	Girod et al.	380/5
5,848,155	12/1998	Cox	380/4
5,850,481	12/1998	Rhoads	382/232
5,859,920	1/1999	Daly et al.	382/115
5,905,819 *	5/1999	Daly	382/284
5,915,027 *	6/1999	Cox et al.	380/54
5,930,369 *	7/1999	Cox et al.	380/54

OTHER PUBLICATIONS

Baritaud, T., et al., "On the Security of the Permuted Kernel Indentification Scheme", Proceedings of the 12th Annual International Cryptology Conference, Advances in Cryptology—Crypto '92, Brickell, E.F., (ed.), Santa Barbara, CA, 305–311 (Aug. 16–20, 1992).

Bender, W., et al., "Techniques for Data Hiding", *IBM Systems Journal, 35*, 313–336 (1996).

Bender, W., et al., "Techniques for Data Hiding", *SPIE, 2420*, 164–173 (1995).

Boland, F.M., et al., "Watermarking Digital Images for Copyright Protection", *IEE International Conference on Image Processing and Its Applications*, Edinburgh, Scotland, 326–330 (Jul. 4–6, 1995).

Boney, L., et al., "Digital Watermarks for Audio Signals", Proceedings of the 1996 IEEE International Conference on Multimedia Computing and Systems, Multimedia '96, Hiroshima, Japan, 473–480 (Jun. 1996).

Bors, A.G., et al., "Image Watermarking Using DCT Domain Constraints", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. III*, Lausanne, Switzerland, 231–234 (Sep. 16–19, 1996).

Bouman, C., et al., "Multiple Resolution Segmentation of Textured Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*, 99–113 (Feb. 1991).

Cawkell, A.E., "Picture–Queries and Picture Databases", *The Journal of Information Science, 19*, 409–423 (1993).

Chalom, E., et al., "Segmentation of an Image Sequence Using Multi–Dimensional Image Attributes", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. II*, Lausanne, Switzerland, 525–528 (Sep. 16–19, 1996).

Chan, W.–Y., et al., "Generalized Product Code Vector Quantization: A Family of Efficient Techniques for Signal Compression", *Digital Signal Processing, 4*, 95–126 (1994).

Chang, S.–F., "Compressed–Domain Techniques for Image/Video Indexing and Manipulatin", *Proceeding fo the 1995 IEEE International Conference on Image Processing, vol. 1*, Washington, D.C., 314–317 (Oct. 23–26, 1995).

Chang, S.–F., et al., "Transform Coding of Arbitrarily-Shaped Image Segments", *Proceedings of the ACM, Multimedia 93*, Anaheim, CA, 83–90 (Aug. 1–6, 1993).

Chitprasert, B., et al., "Human Visual Weighted Progressive Image Transmission", *IEEE Transactions on Communications, 38*, 1040–1044 (Jul. 1990).

Corset, I., et al., "MPEG–4: Very Low Bit Rate Coding for Multimedia Applications", *Proceedings of the SPIE, Visual Communications and Image Processing '94, vol. 2308*, Chicago, IL 1065–1073 (Sep. 25–29, 1994).

Cox, I.J., et al., "Secure Spread Spectrum Watermarking for Images, Audio and Video", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. III*, Lausanne, Switzerland, 243–246 (Sep. 16–19, 1996).

Craver, S., et al., "Can Invisible Watermarks Resolve Rightful Ownership!", *IBM Research Technical Report, RC 20509, IBM CyberJournal*, 23 p. (Jul. 25, 1996).

Daubechies, I., et al., "Orthonormal Bases of Compactly Supported Wavelets", *Communications on Pure and Applied Mathematics, XLI*, 909–996 (Oct. 1988).

Faloutsos, C., et al., "Signature Files: An Access Method for Documents and Its Analytical Performance Evaluation", *ACM Transactions on Office Information Systems, 2*, 267–288 (Oct. 1984).

Flickner, M., et al., "Query by Image and Video Content: The QBIC System", *Computer, 28*, 23–32 (Sep. 1995).

Gary, J.E., et al., "Shape Similarity–Based Retrieval in Image Database Systems", *Proceedings of the SPIE, Image Storage and Retrieval Systems, vol. 1662*, San Jose, CA, 2–8 (Feb. 13–14, 1992).

Girod, B., "The Information Theoretical Significance of Spatial and Temporal Masking in Video Signals", *Proceedings of the SPIE, Human Vision, Visual Processing and Digital Display, vol. 1077*, 178–187 (1989).

Gruber, J., "Smart Paper", *Wired, 2*, 46 (Dec. 1994).

Gudivada, V.N., et al., "Content–Based Image Retrieval Systems", *Computer, 28*, 18–22 (Sep. 1995).

Hartung, F., et al., "Digital Watermarking of Raw and Compressed Video", *SPIE, 2952*, 205–213 (Oct. 1996).

Hirata, K., et al., "Rough Sketch–Based Image Information Retrieval", *NEC Research & Development, 34*, 463–473 (Apr. 1993).

Hirotsugu, K., "An Image Digital Signature System with ZKIP for the Graph Isomorphism", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. III*, Lausanne, Switzerland, 247–250 (Sep. 16–19, 1996).

Hsu, C.–T., et al., "Hidden Signatures in Images", *Proceedings of the 1996 IEEE International Conference on Image Proceedings, vol. III*, Lausanne, Switzerland, 223–226 (Sep. 16–19, 1996).

Huang, Z., et al., "Affine–Invariant B–Spline Moments for Curve Matching", *IEEE Transactions on Image Processing, 5*, 1473–1480 (Oct. 1996).

Huffman, D.A., "A Method for the Construcion of Minimum–Redundancy Codes", *Proceedings of the IRE, 40*, 1098–1101 (1952).

Jacobs, C.E., et al., "Fast Multiresolution Image Querying", *Proceedings of the ACM, SIGGRAPH Conference on Computer Graphics*, Los Angeles, CA, 277–286 (1995).

Jayant, N., et al., "Signal Compression Based on Models of Human Perception", *Proceedings of the IEEE, 81*, 1385–1422 (Oct. 1993).

Johnston, J.D., et al., "Wideband Coding—Perceptual Considerations for Speech and Music", In: *Advances in Speech Signal Processing*, Furui, S., et al., (eds.), Dekker, New York, p. 109–140 (1992).

Le Gall, D., "MPEG: A Video Compression Standard for Multimedia Applications", *Communications of the ACM, 34*, 46–58 (Apr. 1991).

Legge, G.E., et al., "Contrast Masking in Human Vision", *The Journal of the Optical Society of America, 70*, 1458–1471 (Dec. 1980).

Lin, H.–C., et al., "Color Image Retrieval Based on Hidden Markov Models", *Proceedings of the 1995 IEEE International Conference on Image Processing vol. 1*, Washington, D.C., 342–345 (1995).

Macq, B.M., et al., "Cryptology for Digital TV Broadcasting", *Proceeding of the IEEE, 83*, 944–957 (Jun. 1995).

Manjunath, B.S., et al., "Browsing Large Satellite and Aerial Photographs", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. II*, Lausanne, Switzerland, 765–768 (Sep. 16–19, 1996).

Matsui, K., et al., "Video–Steganography: How to Secretly Embed a Signature in a Picture", *IMA Intellectual Property Project Proceedings, vol. 1*, 187–206 (Jan. 1994).

Nam, J., et al., "Combined Audio and Visual Streams Analysis for Video Sequence Segmentation", Proceedings of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing, vol. IV, Munich, Germany, 2665–2668 (Apr. 21–24, 1997).

Niblack, W., et al., "The QBIC Project: Querying Images by Content Using Color, Texture and Shape", *Proceedings of the SPIE, Storage and Retrieval for Image and Video Databases, vol. 1908*, 173–187 (1993).

Nill, N.B., "A Visual Model Weighted Cosine Transform for Image Compression and Quality Assessment", *IEEE Transactions on Communications, COM–33*, 551–557 (Jun. 1985).

Noll, P., "Wideband Speech and Audio Coding", *IEEE Communications Magazine*, 31, 34–44 (Nov. 1993).

O Ruanaidh, J.J.K., et al., "Phase Watermarking of Digital Images", *Proceedings of the 1996 IEEE International Conference on Image Processing vol. III*, Lausanne, Switzerland, 239–242 (Sep. 16–19, 1996).

Pitas, I., "A Method for Signature Casting on Digital Images", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. III*, Lausanne, Switzerland, 215–218 (Sep. 16–19, 1996).

Rioul, O., et al., "Wavelets and Signal Processing", *IEEE Signal Processing Magazine, 8*, 14–38 (Oct. 1991).

Rivest, R.L., "Cryptography", In: *Handbook of Theoretical Computer Sciences*, vol. A, Van Leeuwen, J., (ed.), p. 717–755 (1990).

Rivest, R.L., et al., "A Method for Obtaining Digital Signatures and Public–Key Cryptosystems", *Communications of the ACM 21*, 120–126 (Feb. 1978).

Smith, J.P., "Authentication of Digital Medical Images with Digital Signature Technology", *Radiology, 194*, 771–774 (Mar. 1995).

Smith, J.R., et al., "Modulation and Information Hiding in Images", *Information Hiding*, Proceedings of the First Int. Workshop, Anderson, R., (ed.), Cambridge, U.K., 207–226 (May 30–Jun. 1, 1996).

Srihari, R.K., "Combining Text and Image Information in Content–Based Retrieval", *Proceedings of the 1995 IEEE International Conference on Image Processing*, Washington, D.C., 326–328 (Oct. 23–26, 1995).

Strang, G., "Wavelets and Dilation Equations: A Brief Introduction", *SIAM Review, 31*, 614–627 (Dec. 1989).

Swain, M.J., et al., "Color Indexing", *International Journal of Computer Vision, 7*, 11–32 (1991).

Tanaka, K., et al., "Embedding Secret Information into a Dithered Multi–Level Image", *1990 IEEE Military Communications Conference, vol. 1*, "Milcom 90: A New Era," Monterey, CA, 216–200 (Sep. 30–Oct. 3, 1990).

van Schyndel, R.G., et al., "A Digital Watermark", *Proceedings of the IEEE, ICIP–94, vol. II*, Austin, TX, 86–90 (Nov. 13–16, 1994).

Voyatzis, G., et al., "Applications of Toral Automorphisms in Image Watermarking", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. II*, Lausanne, Switzerland, 237–240 (Sep. 16–19, 1996).

Wallace, G.K., "The JPEG Still Picture Compression Standard", *Communications of the ACM, 34*, 30–44 (Apr. 1991).

Witten, I.H., et al., "Arithmetic Coding for Data Compression", *Communications of the ACM, 30*, 520–540 (Jun. 1987).

Wolfgang, R.B., et al., "A Watermark for Digital Images", *Proceedings of the 1996 IEEE International Conference on Image Processing, vol. III*, Lausanne, Switzerland, 219–222 (Sep. 16–19, 1996).

Wunsch, P., et al., "Wavelet Descriptors for Multiresolution Recognition of Handprinted Characters", *Pattern Recognition, 28*, 1237–1249 (Aug. 1995).

Zhu, B., et al., "Image Coding with Mixed Representations and Visual Masking", *Proceedings of the 1995 IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4*, Detroit, MI, 2327–2330 (May 9–12, 1995).

Zhu, B., et al., "Low Bit Rate Near–Transparent Image Coding", *Proceedings of the SPIE, International Conference on Wavelet Applications for Dual Use, vol. 2491*, Orlando, FL, 173–184. (1995).

Zhu, S.C., et al., "Region Competition: Unifying Snakes, Region Growing, Energy/Bayes/MDL for Multi–band Image Separation", *Proceedings of the IEEE Fifth International Conference on Computer Vision*, Massachusetts Institute of Technology, Cambridge, MA, 416–423 (Jun. 20–23, 1995).

Ziv, J., et al., "A Universal Algorithm for Sequential Data Compression", *IEEE Transactions on Information Theory, IT–23*, 337–343 (May 1977).

* cited by examiner

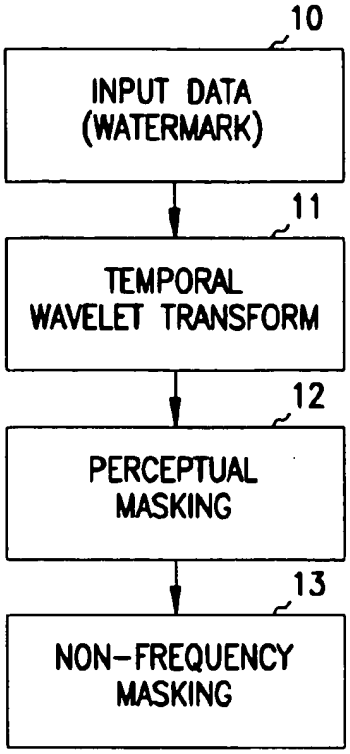

10
INPUT DATA (WATERMARK)
11
TEMPORAL WAVELET TRANSFORM
12
PERCEPTUAL MASKING
13
NON-FREQUENCY MASKING

FIG. 1

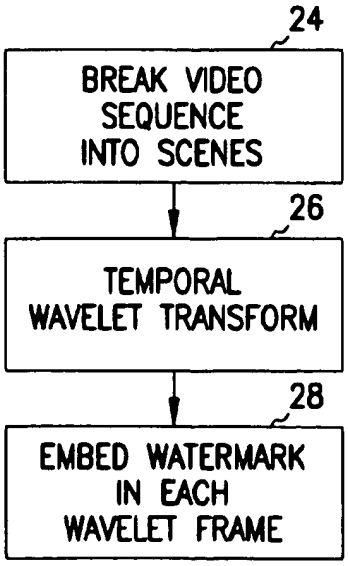

24
BREAK VIDEO SEQUENCE INTO SCENES
26
TEMPORAL WAVELET TRANSFORM
28
EMBED WATERMARK IN EACH WAVELET FRAME

FIG. 2

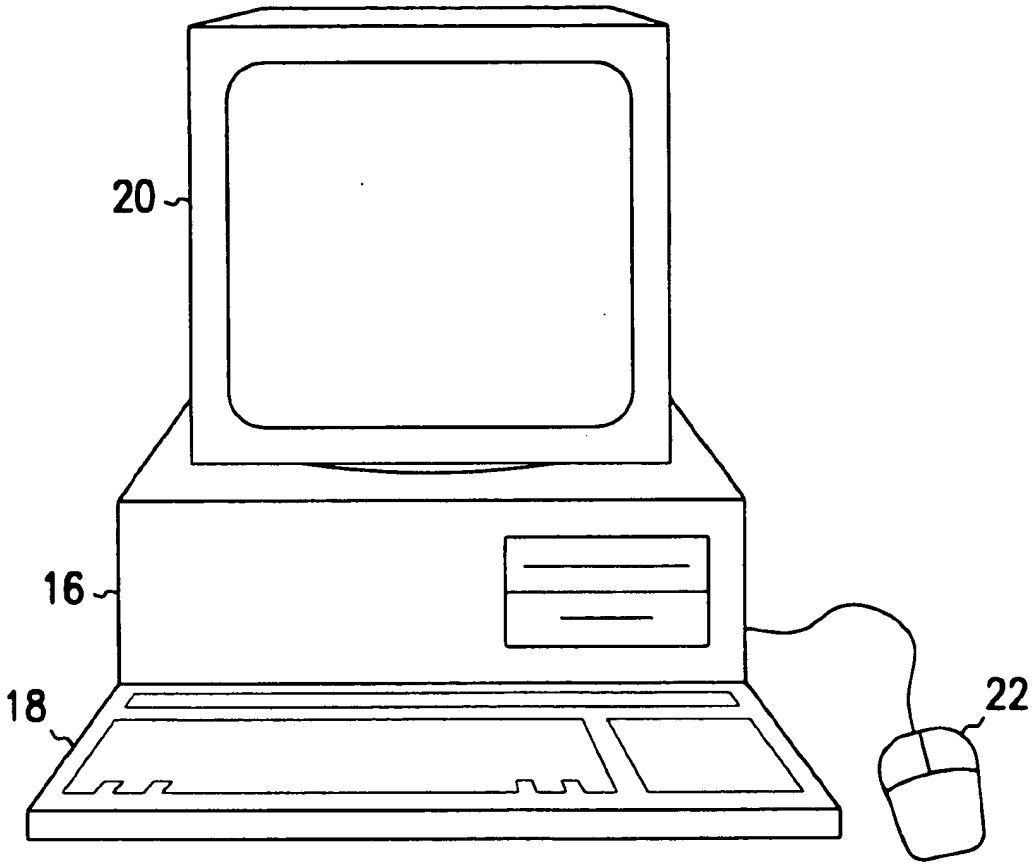


FIG. 3

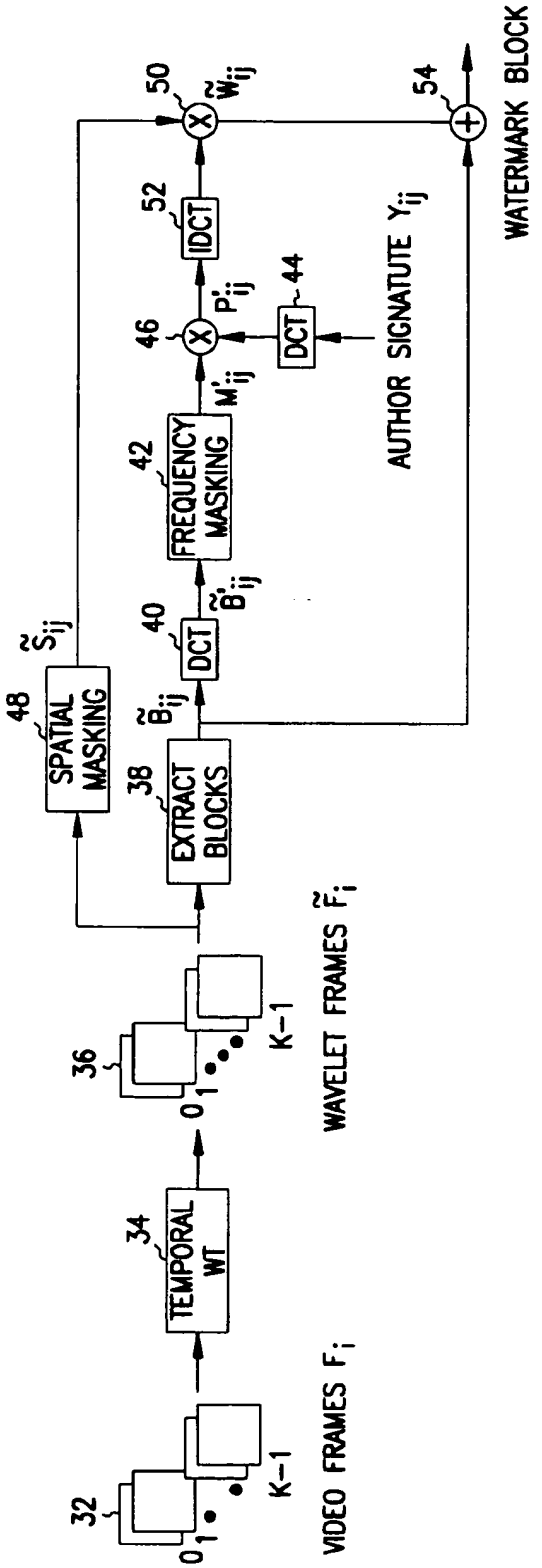

VIDEO FRAMES F_i
32
0 1 K-1
TEMPORAL WT
34
0 1 K-1
36
WAVELET FRAMES $\tilde{F}_i$
EXTRACT BLOCKS
38
$\tilde{B}_{ij}$
DCT
40
$\tilde{B}'_{ij}$
FREQUENCY MASKING
42
M'_{ij}
DCT
44
P'_{ij}
46
IDCT
52
$\tilde{W}_{ij}$
50
AUTHOR SIGNATURE Y_{ij}
54
WATERMARK BLOCK
$\tilde{S}_{ij}$
SPATIAL MASKING
48

FIG. 4

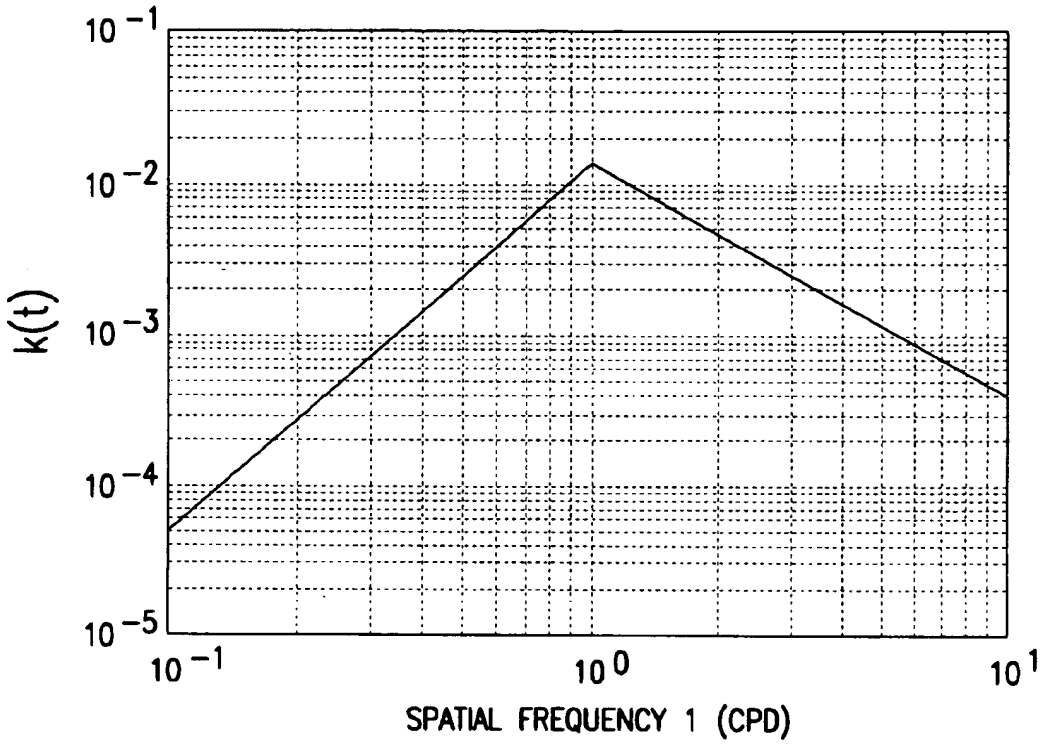


FIG. 5

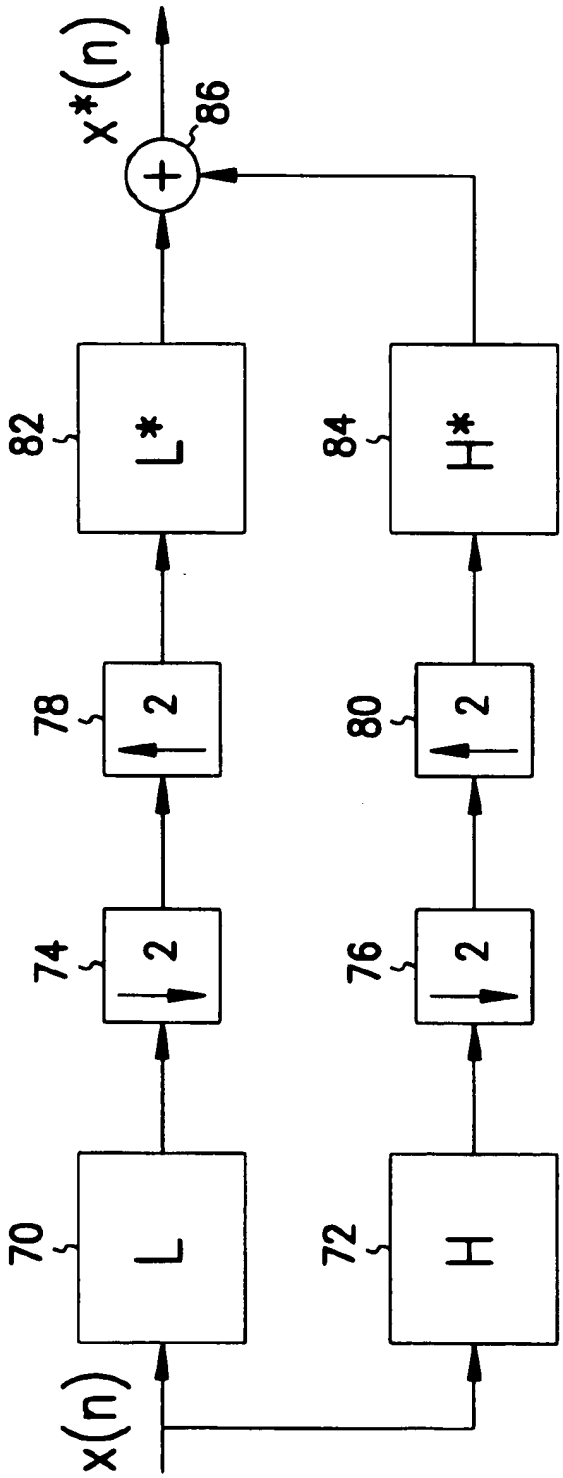


FIG. 6

METHOD AND APPARATUS FOR SCENE-BASED VIDEO WATERMARKING

RELATED DOCUMENTS

This application claims the benefit of U.S. Provisional Application No. 60/024,979, filed Aug. 30, 1996, which is hereby incorporated by reference. U.S. Provisional Application No. 60/050,587, filed Jun. 24, 1997, the benefit of which is also claimed, is also hereby incorporated by reference. Co-filed applications entitled "Method and Apparatus for Embedding Data, Including Watermarks, in Human Perceptible Sounds," Appl. Ser. No. 08/918,891, now U.S. Pat. No. 6,061,793, "Method and Apparatus for Embedding Data, Including Watermarks, in Human Perceptible Images," Appl. Ser. No. 08/918,122, now U.S. Pat. No. 6,031,914 and "Method and Apparatus for Video Watermarking," Appl. Ser. No. 08/918,125, and "Digital Watermarking to Resolve Multiple Claims of Ownership," Appl. Ser. No. 08/918,126 are also hereby incorporated by reference.

STATEMENT REGARDING GOVERNMENT RIGHTS

The present invention was made with government support by AFOSR under grant AF/F49620-94-1-0461, NSF grant INT-9406954, and AF/F49620-93-1-0558. The Government has certain rights in this invention.

FIELD OF THE INVENTION

This invention relates generally to techniques for embedding data such as watermarks, signatures and captions in digital data, and more particularly to scene-based watermarks in digital data that relates to video.

BACKGROUND OF THE INVENTION

Digital video is readily reproduced and distributed over information networks. However, these attractive properties lead to problems enforcing copyright protection. As a result, creators and distributors of digital video are hesitant to provide access to their digital intellectual property. Digital watermarking has been proposed as a means to identify the owner and distribution path of digital data. Digital watermarks address this issue by embedding owner identification directly into the digital data itself. The information is embedded by making small modifications to the pixels in each video frame. When the ownership of a video is in question, the information can be extracted to completely characterize the owner or distributor of the data.

Video watermarking introduces issues that generally do not have a counterpart in images and audio. Video signals are highly redundant by nature, with many frames visually similar to each other. Due to large amounts of data and inherent redundancy between frames, video signals are highly susceptible to pirate attacks, including frame averaging, frame dropping, interpolation, statistical analysis, etc. Many of these attacks may be accomplished with little damage to the video signal. A video watermark must handle such attacks. Furthermore, it should identify any image created from one or more frames in the video.

Furthermore, to be useful, a watermark must be perceptually invisible, statistically undetectable, robust to distortions applied to the host video, and able to resolve multiple ownership claims. Some watermarking techniques modify spatial/temporal data samples, while others modify transform coefficients. A particular problem afflicting all prior art techniques, however, is the resolution of rightful ownership of digital data when multiple ownership claims are made, i.e., the deadlock problem. Watermarking schemes that do not use the original data set to detect the watermark are most vulnerable to deadlock. A pirate simply adds his or her watermark to the watermarked data. It is then impossible to establish who watermarked the data first.

Watermarking procedures that require the original data set for watermark detection also suffer from deadlocks. In such schemes, a party other than the owner may counterfeit a watermark by "subtracting off" a second watermark from the publicly available data and claim the result to be his or her original. This second watermark allows the pirate to claim copyright ownership since he or she can show that both the publicly available data and the original of the rightful owner contain a copy of their counterfeit watermark.

There is a need, therefore, for watermarking procedures applicable to video digital data that do not suffer from the described shortcomings, disadvantages and problems.

SUMMARY OF THE INVENTION

The above-identified shortcomings, disadvantages and problems found within the prior art are addressed by the present invention, which will be understood by reading and studying the following specification. The invention provides for the scene-based watermarking of video data.

In one embodiment of the invention, scenes are extracted from video host data that is made up of a number of successive frames. Each scene thus includes a number of frames. Each frame undergoes a wavelet transformation, which is then segmented into blocks. A frequency mask is applied to the corresponding frequency-domain blocks, which is then weighted with the author signature, also in the frequency domain. The resulting weighted block is taken out of the frequency domain, and then weighted with the spatial mask for its corresponding wavelet transformed block. A unique watermark generation routine is also described that assists in the resolution of deadlock.

The approach of the invention provides advantages over the approaches found in the prior art. In the prior art, an independent watermark applied to each frame may result in detection of the watermark by statistically comparing or averaging similar regions and objects in successive video frames, as has been described in the background. However, the inventive scene-based approach addresses this issue by embedding a watermark this is a composite of static and dynamic components, the dynamic components preventing detection by statistical comparison across frames. Therefore, statistical comparison or averaging does not yield the watermark.

Further aspects, advantages and embodiments of the invention will become apparent by reference to the drawings, and by reading the following detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a flowchart of a method of a video watermarking process according to an embodiment of the invention;

FIG. 2 is a flowchart of a method of an object-based video watermarking process according to an embodiment of the invention;

FIG. 3 is a diagram of a typical computer to be used with embodiments of the invention;

FIG. 4 is a block diagram of a specific implementation of scene-based video watermarking, based on the methods of FIG. 1 and FIG. 2, according to an embodiment of the invention;

FIG. 5 is a diagram showing a masking weighting function k(f) according to one embodiment of the invention; and,

FIG. 6 is a diagram showing a two-band perfect reconstruction filter in accordance with which a wavelet transform can be computed according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

In the following detailed description of the preferred embodiments, reference is made to the accompanying drawings which form a part hereof, and in which is shown by way of illustration specific preferred embodiments in which the invention may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention, and it is to be understood that other embodiments may be utilized and that logical, mechanical and electrical changes may be made without departing from the spirit and scope of the present invention. The following detailed description is, therefore, not to be taken in a limiting sense.

Overview of the Watermarking Process

Referring to FIG. **1**, a flowchart of a method of a video watermarking process, according to one embodiment of the invention, is shown. Specifically, the method of FIG. **1** imbeds watermark data into host video data. In step **10**, the watermark data is generated, which is the signature, or watermark, that acts as a unique identifier for the host video data. Note that the signature inherently is spread across the frequency spectrum without explicit spread-spectrum processing.

In one embodiment of the invention, the signature is a pseudo-random sequence, which is created using a pseudo-random generator and two keys. With the two proper keys, the watermark may be extracted. Without the two keys, the data hidden in the video is statistically invisible and impossible to recover. Pseudo-random generators are well within the art. For example, the reference R. Rivest, "Cryptography," in *Handbook of Theoretical Computer Science* (J. van Leeuwen, ed.), vol. 1, ch. 13, pp. 717–755, Cambridge, Mass.: MIT Press, 1990, which is hereby incorporated by reference, describes such generators.

In one embodiment, the creation of the watermark data in step **10** works as follows. The author has two random keys x1 and x2 (i.e., seeds) from which the pseudo-random sequence y can be generated using a suitable cryptographic operator g(x1,x2), as known within the art. The noise-like sequence y, after some processing, is the actual watermark hidden into the video stream. The key x1 is author dependent. The key x2 is signal dependent. In particular, x1 is the secret key assigned to (or chosen by) the author. Key x2 is computed from the video signal which the author wishes to watermark. The signal dependent key is computed from the masking values of the original signal. The masking values give us tolerable error levels in the host video signal. The tolerable error levels are then hashed to a key x2.

The operator g( ) is called a pseudo-random sequence generator. For the pseudo-random generator to be useful, a pirate must not be able to predict bits of y or infer the keys x1 or x2 from knowledge of some bits of y. There are several popular generators that satisfy these properties, including RSA, Rabin, Blum/Micali, and Blum/Blum/Shub, as known within the art. For example, the Blum/Blum/Shub pseudo-random generator uses the one way function $y=g(x)=x^*x \bmod n$, where $n=pq$ for primes p and q so that $p=q=3 \bmod 4$. It can be shown that generating x or y from partial knowledge of y is computationally infeasible for the Blum/Blum/Shub generator. The classical maximal length pseudo noise sequence (i.e., m-sequence) generated by linear feedback shift registers are not used for this purpose. Sequences generated by shift registers are cryptographically insecure, as one can solve for the feedback pattern (i.e., the keys) given a small number of output bits y.

Thus, a pirate is not free to subtract off a second watermark y' arbitrarily. The pirate must supply the keys x1' and x2' which generate the watermark y' they wish to embed. It is computationally infeasible to invert the one-way function y'=g(x1',x2') to obtain x1' and x2'. Furthermore, x2' is not arbitrary. It is computed directly from the original video signal, which is inaccessible to the pirate. As a result, the two-key pseudo-random sequence author representation resolves the deadlock problem.

In step **11**, a wavelet transform is applied along the temporal axis of the video host data, resulting in a multi-resolution temporal representation of the video. In particular, the representation consists of temporal lowpass frames and highpass frames. The lowpass frames consist of the static components in the video scene. The highpass frames capture the motion components and changing nature of the video sequence (i.e., the video host data). The watermark is designed and embedded in each of these components. The watermarks embedded in the lowpass frames exist throughout the entire video scene. The watermarks embedded in the motion frames are highly localized in time and change rapidly from frame to frame. Thus, the watermark is a composite of static and dynamic components. The combined representation overcomes drawbacks associated with a fixed or independent watermarking procedure. (I.e., avoidance of watermark detection by statistical comparison between successive frames is achieved.)

A wavelet transform can be computed using a two-band perfect reconstruction filter bank as shown in FIG. **6**. The video signal is simultaneous passed through lowpass L filter **70** and highpass H filter **72** and then decimated by 2 (as represented by elements **74** and **76** of FIG. **6**) to give static (no motion) and dynamic (motion) components of the original signal. The two decimated signals may be up sampled (as represented by elements **78** and **80**), and then passed through complementary filters **82** and **84** and summed as represented by element **86** to reconstruct the original signals. Wavelet filters are widely available within the art. For instance, the reference P. P. Vaidyanathan, Multirate Systems and Filter Banks, Englewood Cliffs, N.J.: PTR Prentice-Hall, Inc., 1992, which is hereby incorporated by reference, describes such filters.

Referring back to FIG. **1**, in step **12**, the data generated by step **10** is imbedded into a perceptual mask of the host video data as represented by the temporal wavelet transform of step **11**. The present invention employs perceptual masking models to determine the optimal locations within host data in which to insert the watermark. The perceptual mask is specific to video host data. The mask provides for the watermark data generated by step **10** to be embedded with the host data, at places typically imperceptible to the human eye. That is, the perceptual mask exploits masking properties of the human visual system. Step **12** embeds the watermark within the temporally wavelet transformed host data such that they will not be perceived by a human eye, as defined by the perceptual model. The perceptual masking of step **12** is conducted in the frequency domain.

Thus, image masking models based on the human visual system (HVS) are used to ensure that the watermark embed-

ded into each video frame is perceptually invisible and robust. Visual masking refers to a situation where a signal raises the visual threshold for other signals around it. Masking characteristics are used in high quality low bit rate coding algorithms to further reduce bit rates. The masking models presented here are based on image models.

The masking models give the perceptual tolerance for image coefficients and transform coefficients. These masking models are also described in the reference B. Zhu, et al., "Low Bit Rate Near-Transparent Image Coding," in Proc. of the SPIE Int'l Conf. on Wavelet Apps. for Dual Use, vol. 2491, (Orlando, Fla.), pp. 173–184, 1995, which is hereby incorporated by reference, and in the reference B. Zhu, et al., "Image Coding with Mixed Representations and Visual Masking," in Proc. 1995 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing, (Detroit, Mich.), pp. 2327–2330, 1995, which is also hereby incorporated by reference. The frequency masking model is based on the knowledge that a masking grating raises the visual threshold for signal gratings around the masking frequency. The model is based on the discrete cosine transform (DCT), expresses the contrast threshold at frequency f as a function of f, the masking frequency fm and the masking contrast cm:

$$c(f, f_m)=c_o(f)\cdot\mathrm{Max}\{1, [k(f/f_m)c_m]^\alpha\},$$

where co(f) is the detection threshold at frequency f. The mask weighting function k(f) is shown in FIG. **5**. To find the contrast threshold c(f) at a frequency f in an image, the DCT is first used to transform the image into the frequency domain and find the contrast at each frequency. The value α=0.62 as determined experimentally by psycho-visual tests, and as described in G. E. Legge and J. M. Foley, "Contrast Masking in Human Vision," Journal Optics Society of America, vol. 70, no. 12, pp. 1458–1471 (1980), which is hereby incorporated by reference. Then, a summation rule of the form

$$c(f)=[\Sigma\, f_m\, c(f,f_m)]^{1/\beta},$$

is used to sum up the masking effects from all the masking signals near f. If the contrast error at f is less than c(f), the model predicts that the error is invisible to human eyes.

In step **14**, the host video data as subjected to a temporal wavelet transform in step **11**, with the embedded watermark data from step **12** is further subjected to a non-frequency mask. Because the perceptual mask in step **12** is a frequency domain mask, a further mask is necessary to ensure that the embedded data remains invisible in the host video data. The non-frequency mask is a spatial mask.

Frequency masking effects are localized in the frequency domain, while spatial masking effects are localized in the spatial domain. Spatial masking refers to the situation that an edge raises the perceptual threshold around it. Any model for spatial masking can be used, and such models are well known in the art. However, the model used in one embodiment of the invention is similar to the model described in the Zhu, "Low Bit Rate . . . " reference previously incorporated by referenced, and which is itself based on a model proposed by Girod in "The Information Theoretical Significance of Spatial and Temporal Masking in Video Signals," in Proceedings of the SPIE Human Vision, Visual Processing, and Digital Display, vol. 1077, pp. 178–187 (1989), which is also herein incorporated by reference.

In one embodiment, the upper channel of Girod's model is linearized under the assumption of small perceptual errors, the model giving the tolerable error level for each pixel in the image, as those skilled in the art can appreciate. Furthermore, under certain simplifying assumptions described in the Zhu "Bit Rate . . . " reference, the tolerable error level for a pixel p(x,y) can be obtained by first computing the contrast saturation at (x,y)

$$dc_{sat}(x, y) = dc_{sat} = \sqrt{\frac{T}{\sum_{x',y'} w_4(0, 0, x', y')}}$$

where the weight $w_4(x,y,x',y')$ is a Gaussian centered at the point (x,y) and T is a visual test based threshold. Once $dc_{sat}(x,y)$ is computed, the luminance on the retina, dl_{ret}, is obtained from the equation

$$dc_{sat}(x,y)=w_2(x,y)\cdot dl_{ret}(x,y)$$

From dl_{ret}, the tolerable error level ds(x,y) for the pixel p(x,y) is computed from

$$dl_{ret}(x,y)=w_1(x,y)\cdot ds(x,y)$$

The weights $w_1(x,y)$ and $w_2(x,y)$ are based on Girod's model. The masking model predicts that changes to pixel p(x,y) less than ds(x,y) introduce no perceptible distortion.

As have been described, steps **10**, **11**, **12** and **14** of FIG. **1** provide an overview of the video watermarking process of the present invention. An overview of the scene-based video watermarking process of the present invention is now described.

Overview of the Scene-Based Video Watermarking Process

Referring to FIG. **2**, a flowchart of a method of a scene-based video watermarking process, according to one embodiment of the invention, is shown. The method utilizes the watermarking method of FIG. **1** already described. In step **24**, a video sequence (i.e., the host video data) is broken (segmented) into scenes, as known within the art. For example, the reference J. Nam and A. H. Tewfik, "Combined Audio and Visual Streams Analysis for Video Sequence Segmentation," in Proceedings of the 1997 International Conference on Acoustics, Speech and Signal Processing, (Munich, Germany), pp. 2665–2668 (April 1997), which is hereby incorporated by reference, describes such scene segmentation. Segmentation into scenes allows the watermarking procedures to take into account temporal redundancy. Visually similar regions in the video sequence, e.g., frames from the same scene, must be embedded with a consistent watermark. The invention is not limited to a particular segmentation into scenes algorithm, however.

In step **26**, a temporal wavelet transform is applied on the video scenes, as has been previously described. That is, each scene comprises a number of frames, such that a temporal wavelet transform is applied to each frame within a scene. The resulting frames are known as wavelet frames. The multiresolution nature of the wavelet transform allows the watermark to exist across multiple temporal scales, resolving pirate attacks. For example, the embedded watermark in the lowest frequency (DC) wavelet frame exists in all frames in the scene.

In step **28**, a watermark is embedded in each wavelet frame. The watermark is designed and embedded in the wavelet domain, such that the individual watermarks for each wavelet frame are spread out to varying levels of support in the temporal domain. For example, watermarks embedded in highpass wavelet frames are localized tempo-

rally. Conversely, watermarks embedded in lowpass wavelet frames are generally located throughout the scene in the temporal domain. The watermarks are embedded in accordance with perceptual and non-frequency masks, as has been described. That is, the watermarks are embedded in each frame of each scene in accordance with perceptual and spatial (non-frequency) characteristics of the frame, as has been described in conjunction with the method of FIG. **1**.

The scene-based video watermarking method of the invention has several other advantages. It is scene-based and video dependent, and directly exploits spatial masking, frequency masking, and temporal properties such that the embedded watermark is invisible and robust. The watermark consists of static and dynamic temporal components that are generated from a temporal wavelet transform of the video scenes. The resulting wavelet frames are modified by a perceptually shaped pseudo-random sequence representing the author (owner). The noise-like watermark is statistically undetectable to thwart unauthorized removal. Furthermore, the author representation resolves the deadlock problem. The multiresolution watermark may be detected on single frames without knowledge of the location of the frames in the video scene.

Because the video watermarking procedure is perception-based, the watermark adapts to each individual video signal. In particular, the temporal and frequency distributions of the watermark are controlled by the masking characteristics of the host video signal. As a result, the strength of the watermark increases and decreases with host, e.g., higher amplitude in regions of the video with more textures, edges, and motion. This ensures that the embedded watermark is invisible while having the maximum possible robustness.

Because the watermark representation is scene-based and multiscale, given one or more frames from a potentially pirated video, the watermark may be extracted from the frames without knowledge of the location of the frame being tested. This detection characteristic exists due to the combined static and dynamic representation of the watermark.

The watermark representation of the invention provides an author representation that solves the deadlock problem. The author or owner of the video is represented with a pseudo-random sequence created by a pseudo-random generator and two keys. One key is author dependent, while the second key is signal dependent. The representation is able to resolve rightful ownership in the face of multiple ownership claims.

The watermark representation of the invention also provides a dual watermark. The watermarking scheme uses the original video signal to detect the presence of a watermark. The procedure can handle virtually all types of distortions, including cropping, temporal rescaling, frame dropping, etc., using a generalized likelihood ratio test. This procedure is integrated with a second watermark which does not require the original signal to address the deadlock problem.

As have been described, steps **24**, **26**, and **28** of FIG. **2** provide an overview of the scene-based watermarking process of the present invention. The specifics of the hardware implementation of the invention are now provided.

Hardware Implementation of the Invention

The present invention is not limited as to the type of computer on which it runs. However, a typical example of such a computer is shown in FIG. **3**. Computer **16** is a desktop computer, and may be of any type, including a PC-compatible computer, an Apple Macintosh computer, a UNIX-compatible computer, etc. Computer **16** usually includes keyboard **18**, display device **20** and pointing device **22**. Display device **20** can be any of a number of different devices, including a cathode-ray tube (CRT), etc. Pointing device **22** as shown in FIG. **3** is a mouse, but the invention is not so limited. Not shown is that computer **16** typically also comprises a random-access memory (RAM), a read-only memory (ROM), a central-processing unit (CPU), a fixed storage device such as a hard disk drive, and a removable storage device such as a floppy disk drive. The computer program to implement the present invention is typically written in a language such as C, although the present invention is not so limited.

The specifics of the hardware implementation of the invention have been described. A particular implementation of the scene-based video watermarking of the invention, based on the methods of FIG. **1** and FIG. **2**, is now described.

Particular Implementation of Scene-Based Video Watermarking

The embodiment shown in FIG. **4** illustrates a particular implementation of scene-based video watermarking according to the invention, as based on the methods of FIG. **1** and FIG. **2** that have already been described. Referring now to FIG. **4**, a block diagram of this specific implementation of scene-based video watermarking is shown. Video frames **32** (of video host data) are denoted such that Fi is the ith frame in a video scene, where i=0, . . . , k−1. Frames are ordered sequentially according to time. Each frame is of size n×m. The video itself may be gray scale (8 bits/pixel) or color (24 bits/pixel). Frames **32** undergo a temporal wavelet transformation **34**, as has been described, to become wavelet frames **36**. The tilde representation is used to denote a wavelet representation. For example, F~i is the ith wavelet coefficient frame. Without loss of generality, wavelet frames are ordered from lowest frequency to highest frequency—i.e., F~0 is a DC frame. Thus, there are k wavelet coefficient frames F~i, i=0, . . . , k−1.

In step **38**, each wavelet frame F~i is segmented into 8×8 blocks B~ij, i=0, 1, . . . , (n/8) and j=0, 1, . . . , (m/8). In step **40**, each block B~ij is subjected to a discrete cosine transform (DCT), to become block B~ij'. In step **42**, a perceptual frequency mask, as has been described, is applied to each block to obtain the frequency mask M'ij. In step **44**, author signature Yij—the watermark—also undergoes a discrete cosine transform to become Y'ij. It should be noted that the generation of author signature Yij is desirably in accordance with the process that has been described in conjunction with step **10** of FIG. **1**, but the invention is not so limited.

In step **46**, the mask M'ij is used to weight the noise-like author Y'ij for that frame block, creating the frequency-shaped author signature P'ij=M'ijY'ij. In step **48**, the spatial mask S~ij is generated, as has been described, and in step **50**, the wavelet coefficient watermark block W~ij is obtained by computing the inverse DCT of P'ij in step **52** and locally increasing the watermark to the maximum tolerable error level provided by the spatial mask S~ij. Finally, in step **54**, the watermark W~ij is added to the block B~ij, creating the watermarked block. The process is repeated for each wavelet coefficient frame F~i.

The watermark for each wavelet coefficient frame is the block concatenation of all the watermark blocks for that frame. The wavelet coefficient frames with the embedded watermarks are then converted back to the temporal domain using the inverse wavelet transform. As the watermark is designed and embedded in the wavelet domain, the individual watermarks for each wavelet coefficient frame are

spread out to varying levels of support in the temporal domain. For example, watermarks embedded in highpass wavelet frames are localized temporally. Conversely, watermarks embedded in lowpass wavelet frames are generally located throughout the scene in the temporal domain.

The watermarks embedded within the video data according to the method of FIG. 4 should be extractable even if common signal processing operations are applied to the host data. This is particularly true in the case of deliberate unauthorized attempts to remove the watermark. For example, a pirate may attempt to add noise, filter, code, re-scale, etc., the host data in an attempt to destroy the watermark. The embedded watermark, however, is noise-like and its location over multiplied blocks of the host data, over successive frames of the data, is unknown. Therefore, the pirate has insufficient knowledge to directly remove the watermark. Furthermore, a different signature is used for each block to further reduce unauthorized watermark removal by cross correlation. Any destruction attempts are done blindly.

Detection of the watermark is accomplished via generalized likelihood ratio test. Two methods have been developed to extract the potential watermark from a test video or test video frame. Both employ hypothesis testing. One test employs index knowledge during detection, i.e., the placement of the test video frame(s) relative to the original video is known. The second detection method does not require knowledge of the location of the test frame(s). This is extremely useful in a video setting, where 1000's of frames may be similar, and it is uncertain where the test frames reside.

In the first method, watermark detection with index knowledge, when the location of the test frame is known, a straightforward hypothesis test may applied. For each frame in the test video Rk, a hypothesis test is performed.

H0: Xk=Rk−Fk=Nk (no watermark)

H1: Xk=Rk−Fk=W*k+Nk (watermark)

where Fk is the original frame, W*k is the (potentially modified) watermark recovered from the frame, and Nk is noise. The hypothesis decision is obtained by computing the scalar similarity between each extracted signal and original watermark Wk: Sk=Simk(Xk, Wk)=(Xk*Wk)/(Wk*Wk). The overall similarity between the extracted and original watermark is computed as the mean of Sk for all k: S=mean (Sk). The overall similarity is compared with a threshold to determine whether the test video is watermarked. The experimental threshold is desirably chosen around 0.1, i.e., a similarity value >=0.1 indicates the presence of the owner's copyright. In such a case, the video is deemed the property of the author, and a copyright claim is valid. A similarity value <0.1 indicates the absence of a watermark.

When the length (in terms of frames) of the test video is the same as the length of the original video, the hypothesis test is performed in the wavelet domain. A temporal wavelet transform of the test video is computed to obtain its wavelet coefficient frames R~k. Thus,

H0: X~k=R~k−F~k=Nk (no watermark)

H1: X~k=R~k−F~k=W~*k+Nk (watermark)

where F~k are the wavelet coefficient frames from the original video, W~*k is the potentially modified watermarks from each frame, and Nk is noise. This test is performed for each wavelet frame to obtain X~k for all k. Similarity values are computed as before, Sk=Simk(X~k,W~k).

Using the original video signal to detect the presence of a watermark, virtually all types of distortions can be handled, including cropping, rotation, resealing, etc., by employing a generalized likelihood ratio test. A second detection scheme which is capable of recovering a watermark after many distortions without a generalized likelihood ratio test has also been developed. The procedure is fast and simple, particularly when confronted with the large amount of data associated with video.

In the method for watermark detection without index knowledge, there is no knowledge of the indices of the test frames. Pirate tampering may lead to many types of derived videos which are often difficult to process. For example, a pirate may steal one frame from a video. A pirate may also create a video which is not the same length as the original video. Temporal cropping, frame dropping, and frame interpolation are all examples. A pirate may also swap the order of the frames. Most of the better watermarking schemes currently available use different watermarks for different images. As such, they generally require knowledge of which frame was stolen. If they are unable to ascertain which frame was stolen, they are unable to determine which watermark was used.

This method can extract the watermark without knowledge of where a frame belongs in the video sequence. No information regarding cropping, frame order, interpolated frames, etc., is required. As a result, no searching and correlation computations are required to locate the test frame index. The hypothesis test is formed by removing the low temporal wavelet frame from the test frame and computing the similarity with the watermark for the low temporal wavelet frame. The hypothesis test is formed as

H0: Xk=Rk−F~0=Nk (no watermark)

H1: Xk=Rk−F~0=W~*k+Nk (watermark)

where Rk is the test frame in the spatial domain and F~0 is the lowest temporal wavelet frame. The hypothesis decision is made by computing the scalar similarity between each extracted signal Xk and original watermark for the low temporal wavelet frame W~0: Simk(Xk, W~0). This simple yet powerful approach exploits the wavelet property of varying temporal support.

Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that any arrangement which is calculated to achieve the same purpose may be substituted for the specific embodiments shown. This application is intended to cover any adaptations or variations of the present invention. Therefore, it is manifestly intended that this invention be limited only by the following claims and equivalents thereof.

I claim:

1. A computerized method for embedding data representing a watermark into host data relating to video:

generating the data representing the watermark;

subjecting the host data to a temporal wavelet transform;

embedding the data into the host data, as subjected to the temporal wavelet transform, in accordance with a perceptual mask conducted in the frequency domain; and,

subjecting the host data, including the data embedded therein, to a non-frequency mask.

2. The computerized method of claim 1, wherein the data representing the watermark comprises a pseudo-random sequence.

3. The computerized method of claim 1, wherein generating the data representing the watermark uses a pseudo-random generator and two keys to generate the data.

4. The computerized method of claim 3, wherein the pseudo-random generator is selected from the group comprising RSA, Rabin, Blum/Micali, and Blum/Blum/Shub.

5. The computerized method of claim 1, wherein the perceptual mask comprises a model in which a contrast

threshold at a frequency f is expressed as a function of the frequency f, a masking frequency f_m and a masking contrast c_m,

$$c(f, f_m)=c_o(f)\cdot Max\{1, [k(f/f_m)c_m]^a\},$$

where $c_o(f)$ is a detection threshold at the frequency f.

6. The computerized method of claim 1, wherein the non-frequency mask comprises a spatial mask.

7. The computerized method of claim 1, wherein subjecting the host data to a temporal wavelet transform results in a multiresolution temporal representation of the video having temporal lowpass frames and temporal highpass frames.

8. A scene-based computerized method of watermarking host data relating to video comprising:

segmenting the host data into a plurality of scenes, each scene having a plurality of frames;

subjecting each frame of each scene to a temporal wavelet transform; and,

embedding each frame of each scene, as has been subjected to the temporal wavelet transform, with a watermark in accordance with perceptual and spatial characteristics of the frame.

9. The scene-based computerized method of claim 8, wherein subjecting each frame of each scene to the temporal wavelet transform results in lowpass wavelet frames and highpass wavelet frames.

10. The scene-based computerized method of claim 9, wherein watermarks embedded in lowpass wavelet frames are located throughout the scene in a temporal domain.

11. The scene-based computerized method of claim 9, wherein watermarks embedded in highpass wavelet frames are localized temporally.

12. A computerized system for watermarking host data relating to video and having a plurality of scenes, each scene having a plurality of frames, comprising:

a processor;

a computer-readable medium;

computer-executable instructions executed by the processor from the computer-readable medium comprising:

applying a temporal wavelet transform to each frame;

segmenting each frame of each scene into blocks;

applying a discrete cosine transform (DCT) to each block to generate a frequency block corresponding to the block;

generating a perceptual mask for each frequency block;

applying the DCT to a watermark for each frequency block;

weighting the perceptual mask for each frequency block with the watermark for the frequency block to which the DCT has been applied to generate a frequency-shaped author block;

applying an inverse DCT to each frequency-shaped author block to generate a time-domain block;

generating a spatial mask for each block;

weighting each time-domain block by a spatial mask to generate a watermark block; and,

adding each block to a corresponding watermark block to generate a watermarked block.

13. A computer-readable medium having a computer program stored thereon to cause a suitable equipped computer to perform a method comprising:

applying a temporal wavelet transform to each frame;

segmenting each frame of each scene into blocks;

applying a discrete cosine transform (DCT) to each block to generate a frequency block corresponding to the block;

generating a perceptual mask for each frequency block;

applying the DCT to a watermark for each frequency block;

weighting the perceptual mask for each frequency block with the watermark for the frequency block to which the DCT has been applied to generate a frequency-shaped author block;

applying an inverse DCT to each frequency-shaped author block to generate a time-domain block;

generating a spatial mask for each block;

weighting each time-domain block by a spatial mask to generate a watermark block; and,

adding each block to a corresponding watermark block to generate a watermarked block.

14. The computer-readable medium of claim 13, wherein the computer-readable medium is a floppy disk.

* * * * *